State Per-Capita Income Change Since 1950

Recent Titles in
Contributions in Economics and Economic History

The Sports Stadium as a Municipal Investment
Dean V. Baim

Food and Agrarian Orders in the World-Economy
Philip McMichael

Economic Structures of Antiquity
Morris Silver

Philanthropy and Economic Development
Richard F. America, editor

Public Policy and the Quality of Life: Market Incentives versus Government Planning
Randall G. Holcombe

Studies in Accounting History: Tradition and Innovation for the Twenty-first Century
Atsuo Tsuji and Paul Garner

A New World Order?: Global Transformations in the Late Twentieth Century
David A. Smith and Józef Böröcz, editors

Transforming Russian Enterprises: From State Control to Employee Ownership
John Logue, Sergey Plekhanov, and John Simmons, editors

The Development of Local Public Services, 1650–1860: Lessons from Middletown, Connecticut
Hannah J. McKinney

Rural Development Research: A Foundation for Policy
Thomas D. Rowley, David W. Sears, Glenn L. Nelson, J. Norman Reid, and Marvin J. Yetley, editors

Epistemics of Development Economics: Toward a Methodological Critique and Unity
Kofi Kissi Dompere and Manzur Ejaz

Economic Policy in the Carter Administration
Anthony S. Campagna

State Per-Capita Income Change Since 1950

Sharecropping's Collapse and Other Causes of Convergence

Leonard F. Wheat
and William H. Crown

Contributions in Economics and Economic History,
Number 167

Greenwood Press
Westport, Connecticut • London

Library of Congress Cataloging-in-Publication Data

Wheat, Leonard F.
State per-capita income change since 1950 : sharecropping's collapse and other causes of convergence / Leonard F. Wheat and William H. Crown.
p. cm.—(Contributions in economics and economic history, ISSN 0084–9235 ; no. 167)
Includes bibliographical references and index.
ISBN 0–313–29694–4 (alk. paper)
1. Share cropping—Southern States—History—20th century. 2. Income distribution—Southern States—History—20th century. 3. Income distribution—United States—History—20th century. 4. Southern States—Economic conditions—1945– 5. United States—Economic conditions—1945– I. Crown, William H. II. Title. III. Series.
HD1478.U62A138 1995
339.2'2'097509045—dc20 95–7437

British Library Cataloguing in Publication Data is available.

Copyright © 1995 by Leonard F. Wheat and William H. Crown

All rights reserved. No portion of this book may be reproduced, by any process or technique, without the express written consent of the publisher.

Library of Congress Catalog Card Number: 95–7437
ISBN: 0–313–29694–4
ISSN: 0084–9235

First published in 1995

Greenwood Press, 88 Post Road West, Westport, CT 06881
An imprint of Greenwood Publishing Group, Inc.

Printed in the United States of America

The paper used in this book complies with the Permanent Paper Standard issued by the National Information Standards Organization (Z39.48–1984)

10 9 8 7 6 5 4 3 2 1

CONTENTS

TABLES

PREFACE

This study was inspired by a conspicuous void in the per-capita income literature. State per-capita incomes have been converging for well over a century. To understand why, you must understand what caused state incomes to differ widely in the first place. And that requires recognizing that (*a*) the South has been the nation's low-income region at least since the Civil War, (*b*) the South's low income resulted from conditions traceable to slavery, for example, high black population percentages in southern states, (*c*) the South's most basic slavery-related, poverty-causing condition was the sharecropper–tenant farmer system of agriculture, which replaced slavery, and (*d*) income convergence has been more than anything else a matter of the South's overcoming the conditions that caused its low income.

This summary oversimplifies, of course. Developments affecting nonsouthern regions, particularly the West, have also contributed to per-capita income convergence. And southern developments not related to slavery or to sharecropping have helped boost southern income. Still, an accurate explanation of income convergence must explicitly recognize and emphasize southern conditions and southern changes.

The income literature has not met this requirement. Some of the literature is simply concerned with measuring or verifying income convergence, not with explaining why state and regional incomes converged. Other studies, including some outstanding ones, are not concerned with convergence at all; they deal with the causes of short-run income *divergence* that occurred between 1978 and 1988. The studies that have actually looked at the causes of long-run convergence have sometimes turned to neoclassical growth theory for explanations. The theory is abstract, in some respects ivory towerish, and pays little if any attention to the South, to its poverty-causing conditions, or to how these conditions changed.

A crippling aspect of this abstract theory is that it is designed primarily for explaining growth differences among *countries*. It has little room for causal forces

that operated in the South—racial mix, sharecropping, segregated schools—but are not characteristic of low-income countries. Neoclassical theory also suffers from other defects—defects exposed by countertheoretical migration patterns and manufacturing growth patterns in the United States since 1950.

One widely publicized and generally well-received study develops a new theory, the institutional sclerosis theory. This theory is similarly abstract, shows little awareness that low income has been a predominantly southern problem, shows even less awareness of the conditions that caused the South's income to be low, and shows no awareness of critical developments in southern agriculture, southern educational levels, and black migration. Moreover, the theory emphasizes differences in manufacturing growth rates (supposedly controlled by differences in degrees of sclerosis) as the immediate cause of convergence. This idea does not square with the fact that the U.S. West, which had the highest per-capita income in 1950, has had the highest manufacturing growth rates but the lowest per-capita income growth rates. The theory implies that the West should have *diverged*. And as the region with the lowest per-capita manufacturing employment in 1950, it should have had the lowest per-capita income—certainly not the highest.

There have been, to be sure, a few excellent down-to-earth studies that have analyzed the causes of long-run income convergence from a nontheoretical perspective. But these studies have a different problem. They have used national income accounting approaches that disaggregate income into its major components. The decompositional approach is ill-suited for explanations that do not involve changes in per-capita employment or industry mix. These studies do provide valuable insights, and they do pay attention to real-world developments that influenced the accounting figures they discuss. But we prefer a different perspective, one with heavier emphasis on socioeconomic causation and on recent economic history.

A broader problem with the income literature—we have already touched on this—is that none of it spotlights or even dimly illuminates what we regard as the most fundamental cause of income convergence since 1950: the collapse of the sharecropper-tenant system in the South. To be sure, many income studies have mentioned agriculture, including southern agriculture; farm-to-urban population shifts have not been ignored. But we are not aware of any *income* literature that even mentions sharecropping. Much less have we seen any understanding of sharecropping's powerful contribution to southern poverty or of the multifaceted effect of sharecropping's post-1950 collapse.

This is not to say that we are the discoverers of sharecropping. Over the decades, quite a number of books and journal articles have discussed the topic. Two fairly recent books—one by Gavin Wright, one by Nicholas Lemann—provide thorough, competent discussions of sharecropping's downfall. But none of this literature is *income* literature; it deals with the southern economy, migration, agriculture, politics, and other topics. In fact, most of this literature doesn't even mention per-capita income. Meanwhile, the previously mentioned studies that *have* dealt with income do not seem to be aware of sharecropping.

Another general problem with the income literature is that none of it tries to explain why per-capita income growth has been slowest in the West, formerly the

high-income region. Neither has the literature recognized or tried to resolve the contradictions (*a*) between low industrialization and high per-capita income in the West in 1950 and (*b*) between high industrial growth and low per-capita income growth in the West since 1950.

All of these considerations convinced us that a new study with a new approach was necessary. This study is the result. Our research was supported by a research grant to Brandeis University by the Economic Development Administration of the U.S. Department of Commerce.

State Per-Capita Income Change Since 1950

1

PER-CAPITA INCOME CONVERGENCE AND DIVERGENCE

State and regional per-capita incomes in the United States have been converging throughout this century.[1] Indeed, the convergence trend reaches back at least as far as 1880[2] and apparently to 1840,[3] if not earlier. The convergence since 1929, the first year of the government's income time series, is especially well documented.[4] Temporary backsliding did occur during the 1920s[5] and 1980s.[6] But over the long run, the below-average states have been climbing toward the national average; the above-average states have been descending. Moreover, incomes have converged within as well as among regions.[7]

Several subtrends are conspicuous. The dominant subtrend has been a strong upward movement in southern per-capita income: the South, which has always stood apart from the other regions as *the* low-income region, has been catching up. Another important subtrend has been relatively slow per-capita income growth in the West, which had the nation's highest per-capita income in 1950. This western subtrend flies in the face of the well-known fact that the West has had the nation's fastest growth by just about any measure of percentage growth in size—total income, population, net in-migration, total employment, or manufacturing employment. A third subtrend has been comparatively fast income growth in the Plains (West North Central) vis-à-vis the Manufacturing Belt and the West.

Looking at state per-capita income changes since 1950, this study seeks to determine the causes of the general trend (convergence) and the subtrends. The econometric part of our analysis deals with 1950–87 change. The 1950 starting point rules out the strong convergence-producing effects of World War II and hard-to-evaluate effects of the immediate postwar readjustment years. The precise date, 1950, was chosen for census reasons. The 1987 termination date was dictated by the data available when the study began. We have, however, been able to carry our trend analysis forward to 1993.

1950-1987 INCOME TRENDS

To understand per-capita income change, one must first understand that high and low per-capita income are not randomly distributed geographically. Per-capita income variation is essentially a regional phenomenon. The states within particular regions tend to be fairly consistently high, consistently low, or consistently somewhere in between. This tendency occurs because states within regions tend to be similar in terms of things that influence income—agriculture, urbanization, manufacturing, education, racial mix, per-capita employment, wage levels, and shipping distance (transport cost) to the region from the Manufacturing Belt. Some of these things, in turn, depend on a state's status as a former slave or nonslave state; this is why the South stands out from the other regions in per-capita income and many related respects. As regional characteristics change, and to the extent that they change, per-capita income changes. Understanding per-capita income change therefore requires understanding how regions are changing in their characteristics.

Table 1.1 provides a basis for such an understanding. It shows the 1950 and 1987 per-capita incomes for the United States, the combined nonsouthern regions, four specific regions, and seven subregions. Per-capita income is given both in 1987 dollars and as a percentage of the national average. The table also shows the 1950–87 percentage increase in per-capita income for each geographic entity. The regions comprise the 48 contiguous states and the District of Columbia. (In the study's econometric models, however, the District is excluded. It is a legitimate part of a region, but it is not substantively a state—it is a city—and should not be used as a state for cross-sectional analysis purposes.) The regional configuration used here is based on that of the Commerce Department's Bureau of Economic Analysis (BEA) but differs in three respects. First, it groups New England, the Mideast, and the Great Lakes into the Manufacturing Belt. Second, it splits BEA's Southwest region, putting Oklahoma and Texas in the South and putting New Mexico and Arizona in the Rocky Mountain subregion. Third, it redivides the South into the Peripheral South (Virginia, West Virginia, Kentucky, Oklahoma, Texas, and Florida) and the Deep South (North Carolina, South Carolina, Georgia, Tennessee, Alabama, Mississippi, Arkansas, and Louisiana). All of these changes help create regions whose characteristics are better related to per-capita income and its determinants.

1950 Per-Capita Income

In 1950, all major regions except the South were either above the national average or (in the case of the Plains) within 3 percent of it. But the South was 26 percent below the national average. The national average incorporates southern data, so a better standard of comparison is the per-capita income of the three nonsouthern regions combined. The South's 1950 per-capita income was less than two-thirds (66 percent) of nonsouthern per-capita income. In the Deep South, per-capita income was only 59 percent of the non-South's; the Peripheral South reached only 74 percent of the non-South level. The West's high per-capita

Table 1.1
1950 and 1987 Per-Capita Income and 1950–1987 Percentage Increase for the United States, Four Regions, and Seven Subregions (48 States plus Washington, D.C.)

	Per-Capita Income (1987 Dollars)		Percentage of National Average		Percentage Increase
	1950	1987	1950	1987	1950–87
United States (49 Units)	$7,059	$15,465			119
Nonsouthern States	7,820	16,375	111	106	109
Regions					
Manufacturing Belt	7,937	16,733	112	108	111
South	5,197	13,556	74	88	161
Plains (W. North Central)	6,848	14,669	97	95	114
West	8,079	16,269	114	105	101
Subregions					
Manufacturing Belt					
New England (ME, NH, VT, MA, RI, CT)	7,539	18,626	107	120	147
Mideast (NY, NJ, PA, DE, MD, DC)	8,155	17,629	116	114	116
Great Lakes (OH, IN, IL, MI, WI)	7,820	15,229	111	98	95
South					
Peripheral South (VA, WV, KY, OK, TX, FL)	5,788	14,316	82	93	147
Deep South (NC, SC, GA, TN, AL, MS, AR, LA)	4,641	12,594	66	81	171
West					
Rocky Mountain (MT, WY, ID, CO, UT, NM, AZ)	6,724	13,609	95	88	102
Far West (WA, OR, NV, CA)	8,532	17,169	121	111	101

Source: Computed from time-series data provided by the Bureau of Economic Analysis, U.S. Department of Commerce.

income rested entirely on the Far West subregion, which was 21 percent above the national average. The West's Rocky Mountain subregion stood at 95 percent of the national average.

The 1950–87 changes will be easier to understand if we quickly review the main reasons for the 1950 differences. The South's singular position among the regions resulted from a set of interrelated characteristics traceable to slavery. Prominent among these southern characteristics were (*a*) the huge poverty population associated with the sharecropper-tenant farmer system of agriculture, (*b*) the high black population percentage, (*c*) low educational attainment levels, (*d*) low proportions of urban and metropolitan population, (*e*) low per-capita employment, and (*f*) low hourly wage rates. These characteristics were strongest in the Deep South, which has the "black belt" states, so that is where the very lowest income was.

Per-capita income was highest in the West partly because the West was relatively urbanized; it had a slightly higher urban population percentage than the East North Central (Great Lakes) region. Urban places enjoy a more favorable mix of industries and occupations. People tend to think of the West as being rural and agricultural, but this is not generally the case. Much of the West is too arid and mountainous for extensive agriculture; the predominant form of agriculture in much of the Rocky Mountain subregion is sheep and cattle ranching. Another reason for high income in the West was its high wage and salary scales. As in Alaska today, the high wages and salaries were needed to accommodate the region's (especially the Far West's) high cost of living. This high cost of living resulted from high transportation costs on manufactured goods imported from the Manufacturing Belt.

The Manufacturing Belt, despite being more industrialized than the West, had somewhat lower per-capita income. The reasons were a slightly lower urban percentage than the West's, a lower cost of living (lower transport costs on manufactured goods), lower educational attainment, and a higher black population percentage. High population percentages of foreign immigrants and Puerto Ricans may also have weakened Manufacturing Belt income.

The Plains ranked third among the nonsouthern regions because it was more agricultural, less urbanized. And largely because it *was* more agricultural, the Plains had a lower median for years of schooling than the other nonsouthern regions.

1950-1987 Percentage Increase

The 1950–87 per capita income change picture is one of convergence. Table 1.1 shows that the South's percentage increase of 161 percent was far above both the 119 percent national average increase and the 109 percent nonsouthern increase. All three nonsouthern regions had per-capita income increases of less than the national average. More striking is the perfect negative relationship between 1950 per-capita income rank and 1950–87 percentage increase rank. The two ranks are as follows:

Rank	1950 Per-Capita Income	1950-87 Percentage Increase
1	West	South
2	Manufacturing Belt	Plains
3	Plains	Manufacturing Belt
4	South	West

At the subregional level, the subregion with the very lowest 1950 per-capita income (Deep South) had the very highest growth (171 percent). The subregion with the very highest 1950 income (Far West) had the lowest growth (101 percent). Again, the result was convergence.

Why did incomes converge? Developments in all four regions contributed. Later in this study these developments will be described in detail; right now we will merely summarize the main themes. Southern developments are foremost. Easily the most important convergence-producing factor was the collapse of the sharecropper-tenant farmer system in the South. This collapse (*a*) shifted poverty to other regions by causing millions of black and poor white sharecroppers and tenants to migrate from the South, (*b*) contributed heavily to southern urbanization by also causing farm-to-urban migration within the South, (*c*) gave the South a more capital-intensive, higher-income form of agriculture, (*d*) raised the South's educational level, (*e*) helped raise per-capita employment, which is higher in urban areas, and (*f*) led to higher hourly wages by eliminating the source of the South's surplus labor. The civil rights movement smashed educational and employment barriers, further enhancing southern educational attainment (better-educated persons earn more) and raising per-capita employment. Rapid manufacturing growth in the South directly created a more favorable industrial mix; further stimulated urbanization, with its richer mix of professional and technical employment; and soaked up surplus labor, causing wages to rise.

The West had the next most important development. Something happened that caused it to slip from first place in per-capita income (ahead of the Manufacturing Belt) to second—even though the Manufacturing Belt had to absorb most of the sharecroppers who left the South. What happened was that the transport cost element in the price of manufactured goods sold in the West fell sharply. Transportation improvements affecting the cost of shipping goods to the West from the Manufacturing Belt helped; so did extremely fast western manufacturing growth, which lessened the West's dependence on eastern goods, with their high shipping costs. These developments lowered the West's cost of living. Wages and salaries responded by falling relative to national averages. The wage-salary decline caused a commensurate decline in relative income.

In the Plains (West North Central) states, income growth was second fastest. The so-called technological revolution in agriculture was largely responsible. New agricultural technology had the most labor-saving effect in the South. There the mechanical cotton picker, late proliferation of tractor technology, and the substitution of chemical weeding (herbicides) for sharecropper hoeing brought unusually big labor savings. But developments in farm equipment, fertilizing, and plant genetics brought gains in other regions too. Across the nation, surplus farmers moved to urban areas and took higher-paying urban jobs. As the nation's second-most agricultural region (behind the South), the Plains saw relatively large

farm-to-urban population shifts—hence relatively large income gains. At the same time, the Plains was comparatively insulated from other developments that adversely affected income in the West and in the Manufacturing Belt.

Industrial self-sufficiency protected the Manufacturing Belt from the transport cost developments that restrained the West's income growth. Yet the Manufacturing Belt still had the second-slowest income growth. The main reason was surely the South-to-North migration of millions of black and poor white sharecroppers and tenant farmers. Relatively few of the migrants went to the Plains or the West; the Manufacturing Belt bore the brunt of the interregional shift in poverty. Lack of agriculture further weakened the region's income growth. Being the nation's least agricultural region, the Manufacturing Belt had the smallest farm-to-urban population shift.

The slow income growth in the Manufacturing Belt was largely the result of slow growth in the Great Lakes subregion; New England and the Mideast actually grew faster than the Plains. The Great Lakes got the lion's share of the Manufacturing Belt's sharecropper in-migrants. This happened because the Great Lakes was closest to the Lower Mississippi Valley, which had the South's heaviest concentration of sharecropping. Income growth in the Great Lakes subregion was also hampered by hard times in the automobile, steel, and rubber industries. The emergence of strong foreign competition in automobiles and steel helped bring on the hard times.

A purely statistical effect also helped cause income convergence in the West and the Manufacturing Belt. The South's fast income growth raised the national average faster than income was growing in the three highest regions. Just as southern income was catching up with the national average, the national average was catching up with western and Manufacturing Belt income. In other words, even in the absence of any negative influences in the West and the Manufacturing Belt, the South's fast income growth would have caused *relative* decline in the high-income regions.

This same statistical effect caused income divergence in the Plains and the Rocky Mountain states. The Plains went from 97 percent of the national average (1950) to 95 percent (1987) not because something went wrong in the Plains but because the region's 114 percent growth rate was less than the national average growth rate of 119 percent; the national average rate embodied the South's 161 percent growth rate. Similarly, the Rocky Mountain subregion went from 95 percent of the national average to 88 percent partly because fast income growth in the South boosted the national average. Rocky Mountain income was also affected by the West's cost-of-living decline.

1987 Per-Capita Income

As a result of the 1950–87 convergence, per-capita income showed far less regional inequality in 1987 than in 1950. Per-capita income in the South and its two subregions equaled the following percentages of *nonsouthern* per-capita income in the two years:

Region or Subregion	1950	1987
South	66%	83%
Deep South	59	77
Peripheral South	74	87

Note that, in all three comparisons, the income gap narrows in 1987 to about (or else exactly) half of its 1950 size. The gap (100 percent minus regional percentage) goes from 34 to 17 percent for the overall South, from 41 to 23 percent for the Deep South, and from 26 to 13 percent for the Peripheral South. For the Peripheral South, the 1950-87 gain is enough to put the subregion ahead of the Rocky Mountain subregion and almost on par with the Plains.

Another way to compare 1950 and 1987 inequality is to look at the percentage point spread between the highest and lowest regions in each of the two years. When income is expressed as a percentage of the national average, the spread (highest region minus lowest) goes from 40 percentage points in 1950 to 20 points in 1987. Measured this way, inequality again declined to half its original amount between 1950 and 1987.

Regional per-capita income as a percentage of the national average decreased in each of the three nonsouthern regions—from 112 to 108 percent in the Manufacturing Belt, from 97 to 95 percent in the Plains, and from 114 to 105 percent in the West. A major part of the explanation in each case is the statistical effect described earlier: big income increases in the South pushed the national average up faster than income in the nonsouthern regions was growing. The extra-big decline in the West seems to be the result of the relative wage and salary declines that resulted from transport-cost-based cost of living declines.

New England and the Rocky Mountain states bucked the nonsouthern trend toward the national average. New England's per-capita income went from 107 percent of the national average in 1950 to 120 percent in 1987. Although five of the six New England states gained relative to the average, most of the gain was in Massachusetts and New Hampshire. Massachusetts went from 98 to 124 percent of the average, New Hampshire from 90 to 116 percent. These two increases occurred mostly during the relatively short-lived New England economic boom of the 1980s; defense contracts are believed to have played an important role. The Rocky Mountain states went from 95 percent of the national average in 1950 to 88 percent in 1987. Relative income declined in the Rockies for the same reason it declined in the overall West. The difference is that the overall West was well above average in 1950, whereas the Rocky Mountain subregion was below average: decline caused convergence for the overall West but divergence for the Rockies.

THE BROADER PICTURE: TEMPORARY DIVERGENCE AND RECONVERGENCE

Our 1950-87 study period falls within a broader period, 1929-93, for which data are now available. Income convergence has not been continuous. Inequality

reached a trough in 1978, then increased over the next ten years to a 1988 peak (at a mid-1960s inequality level). Thereafter convergence resumed. In every year after 1988 (through 1993, the latest year with data), state income inequality was less than in the preceding year. In 1993, income inequality was at about the 1972 level, or the level that prevailed just before the 1973 energy crisis (Arab oil embargo) temporarily boosted income in many oil-producing and coal-producing states—states with generally low income.

We use the coefficient of variation (CV) to measure trends in inequality and to establish the 1978 and 1988 turning points. CV is the standard deviation (SD) of a distribution divided by the mean (M): CV = SD/M. In effect, CV for state per-capita income is SD put on a constant-dollar basis. SD measures the amount of dispersion, or the general magnitude of deviation of the values in a distribution from their mean. Basically, SD is a weighted-average deviation for the cases being averaged. The deviations are first squared to give the bigger deviations more weight. The average squared deviation is then unsquared (converted to its square root) to restore the original order of magnitude.

SD is not appropriate for comparing changes in equality over time, because increases in the mean typically cause increases in SD. Specifically, income grows over time—because both prices (the cost of living) and real income rise. And as income rises, the range of state incomes—hence SD—increases. Time comparisons therefore require that the effect of secular increases in the mean be controlled. That is, time comparisons call for a measure of *relative* dispersion (CV) rather than of absolute dispersion (SD). Dividing SD by M to get CV provides the measure of relative dispersion. The larger CV is, the more dispersion or inequality there is.

Table 1.2 shows CV for the years from 1929 to 1993. The basic picture is this: CV declined from 0.356 in 1929 to 0.216 in 1950 and 0.124 in 1978, then rose to 0.173 in 1988 before resuming a steady downward trend. CV was 0.142 in 1993.

The biggest CV decline for any decade was in the 1940s. This decline occurred because (*a*) World War II boosted southern income in several ways, (*b*) the exodus of sharecroppers from southern farms began to speed up, and (*c*) worldwide food shortages raised income in agricultural states generally. Smaller but still impressive decreases in inequality during the 1950s and 1960s largely reflect the direct and indirect effects of sharecropping's collapse. These decreases also reflect the nationwide effects of the technological revolution in agriculture (farm-to-urban population shifts), rural state (especially southern) educational gains, rail and highway transportation improvements (affecting the West's price-wage structure), and other influences.

Most of the sharecroppers had left the southern farms by 1970, so the mild acceleration in CV's rate of decline over the first eight years of the 1970s seems to have a different explanation. The energy crises of 1973 (OPEC oil embargo) and 1979 (Iranian revolution) raised income in the oil-producing and coal-producing states, most of which were below the national average in income. Texas and Louisiana, ranked one-two in crude oil production in the 1970s, illustrate this development. Texas's income went from 89 percent of the national

Table 1.2
Coefficient of Variation for Per-Capita Income: Forty-eight States, 1929–1993

Year	CV	Year	CV	Year	CV	Year	CV
1929	0.356						
1930	0.379	1950	0.216	1970	0.156	1990	0.161
1931	0.388	1951	0.210	1971	0.153	1991	0.154
1932	0.406	1952	0.206	1972	0.143	1992	0.151
1933	0.393	1953	0.205	1973	0.134	1993	0.142
1934	0.371	1954	0.202	1974	0.132		
1935	0.342	1955	0.202	1975	0.133		
1936	0.352	1956	0.201	1976	0.130		
1937	0.337	1957	0.200	1977	0.127		
1938	0.343	1958	0.184	1978	0.124		
1939	0.348	1959	0.188	1979	0.125		
1940	0.351	1960	0.192	1980	0.131		
1941	0.324	1961	0.190	1981	0.128		
1942	0.285	1962	0.184	1982	0.136		
1943	0.272	1963	0.182	1983	0.140		
1944	0.248	1964	0.184	1984	0.145		
1945	0.238	1965	0.170	1985	0.153		
1946	0.245	1966	0.160	1986	0.161		
1947	0.233	1967	0.158	1987	0.168		
1948	0.214	1968	0.158	1988	0.173		
1949	0.213	1969	0.162	1989	0.169		

Sources: Our 48-state CVs for 1929, 1940, 1950, 1960, 1970, 1978–80, and 1987–93 were calculated directly from the Bureau of Economic Analysis (BEA) state per-capita income time series. CVs for the other years between 1929 and 1970 were estimated from (*a*) 49-state (including D.C.) annual CVs from Amos (1989), (*b*) the differences between our 48-state CVs and the Amos 49-state CVs for our benchmark years, and (*c*) the average annual rates of change in these differences between benchmarks. CVs for 1981–86 were estimated the same way from annual 50-state CVs prepared by Rowley, Redman, and Angle (1991) and our own 1980 and 1987 48-state CVs, used as benchmarks for measuring differences (48 states versus 50). The CVs for 1971–77 are averages of the two estimates, one based on Amos (49 states) and one based on Rowley, Redman, and Angle (50 states).

average in 1972 to 98 percent in 1979; Louisiana's went from 76 percent in 1972 to 85 percent in 1979. Also, the fruits of the southern educational gains initiated by the civil rights movement of the 1960s were becoming more visible by the 1970s.

Inequality increased between 1978 and 1988. In chapter 2 we examine the reasons for the 1978–88 reversal of trend. Briefly, the reversal was caused by a temporary farm crisis (hurting low-income rural states), a severe post-1981 decline in energy prices (also hurting low-income states), and a temporary economic boom in the Northeast (helping high-income states). Most of the increase in inequality occurred after 1981, the year when both energy prices and agricultural prices began to decline. Between 1981 and 1988, income in Iowa (a farm state) went from 99 percent of the national average to 89 percent, income in Texas (an oil state) went from 102 percent to 88 percent, and income in Massachusetts (a northeastern state) went from 108 percent to 126 percent. The three adverse developments played themselves out in the late 1980s, and the long-run influences regained control. CV resumed its decline.

Although 1987–93 is not part of our 1950–87 study period, the recent decrease in inequality is highly relevant to our analysis. The steady downward trend since 1988 establishes that the 1978–88 increase in CV was an aberration and that the long-run downward trend has resumed. CV's 1993 level, 0.142, is better than the 1972 level of 0.143 but still above the 1978 trough level of 0.124.

Why did 1993 inequality still exceed 1978 (trough) inequality? The main reason seems to be the return of the *relative* prices of energy commodities to pre-energy-crisis levels—that is, to the 1972 relationship between energy prices and general prices. Energy prices rebounded only slightly after their long decline during the 1980s. At the end of 1993 the price index for motor fuel stood at 68 percent of the consumer price index (CPI); the price index for fuel oil and other household fuel commodities was 84 percent of CPI. (All three price indexes have 100 as their 1982–1983–1984 average.) The two 1993 energy price percentages were about the same as the percentages found in the early 1970s: the motor fuel price index was 68 percent of CPI in 1973 (compared to 68 percent in 1993); the fuel oil and other household fuel index was 78 percent in 1972 (compared to 84 percent in 1993).[8] The energy states—mostly low-income states—had thus lost by 1993 the income gains they got from the 1973–81 energy price increases. Texas's income, for example, was 89 percent of the national average in 1972 and 92 percent in 1993. Louisiana's relative income was 76 percent in 1972 and 80 percent in 1993.

Another reason why 1993 inequality exceeded 1978 inequality concerns New England, now the subregion with the highest income. New England's 1993 income was 116 percent of the U.S. average. This level was 13 percentage points higher than New England's 1978 level. Connecticut was at 135 percent of the national average in 1993 compared to 117 percent in 1978. Massachusetts was at 118 percent of the national average in 1993 compared to 104 percent in 1978. The reasons for the persistence of high income levels in New England are poorly understood but seem to involve the emergence of high-tech manufacturing.[9]

OVERVIEW OF STUDY

The rest of this study moves toward and ultimately reaches conclusions about the significance and relative importance of the theoretical causes of per-capita income convergence in the United States. Chapter 2 looks at the empirical literature bearing on the reasons for income change. Chapter 3 discusses theories, including our own, that attempt to explain income change. Chapter 4 amplifies the explanations for convergence given in chapters 1 and 3. The result is an organized set of testable hypotheses. Chapter 5 describes the study's correlation-regression methodology and introduces the study variables. The methodology employs two general models, each of which has many variations (equations of varying length and content). The two basic models differ in the way their dependent and independent variables measure change. Both models examine the 48 contiguous states. Chapter 6 presents the regression findings for the two models. Chapter 7 pulls the findings together to assess the individual hypotheses and rank the causes of income change by importance.

To avoid wordiness and undue repetition, we will usually refer to per-capita income simply as "income." Unless the context clearly shows we are referring to total (absolute) income, "income" should be interpreted as meaning *per-capita* income.

NOTES

1. Harvey S. Perloff, Edgar S. Dunn, Jr., Eric E. Lampard, and Richard F. Muth, *Regions, Resources and Economic Growth* (Baltimore: Johns Hopkins University Press, 1960), 187; Gavin Wright, *Old South, New South* (New York: Basic Books, 1986), 55, 240; Benjamin Chinitz, "The Regional Transformation of the American Economy," *Urban Studies* 23 (1986), 377–85; and Robert J. Barro and Xavier Sala-i-Martin, "Convergence," *Journal of Political Economy* 100 (1992), 223–51.

2. Perloff et al., *Regions*, 187, and Barro and Sala-i-Martin, "Convergence."

3. Barro and Sala-i-Martin, "Convergence." The authors state: "In 1840, the southern and nonsouthern states differed little in terms of average per capita income: the unweighted average of 11 southern states was 94 percent of that for 18 eastern and midwestern states" (p. 236). But we would interpret their data (figure 2, p. 236) differently. In 1840 the four lowest per-capita incomes belonged to northern frontier states: Iowa (not yet a state), Michigan (1837), Illinois (1818), and Indiana (1816). Two other northern states—Maryland and Delaware—were slave states that could well be classified as southern, which is how the Census Bureau classifies them. If we eliminate from the 29 states and states-to-be the newest 10 (admitted after 1812) and the two "northern" slave states, 17 states are left. Now 8 of the 9 states with the highest per-capita income (all but Louisiana) are northern, and 6 of the 8 lowest states (all but Ohio and Maine) are southern. Thus, even in 1840, South-non-South inequality and the depressing effects of slavery on income were conspicuous.

4. Lynn E. Browne, "Narrowing Regional Income Differentials," *New England Economic Review* (September-October 1980) 35–54; Daniel H. Garnick and Howard L. Friedenberg, "Accounting for Regional Differences in Per Capita Income Growth, 1929–79," *Survey of Current Business* 62 (September 1982), 24–34; Cletus C. Coughlin

and Thomas B. Mandelbaum, "Why Have State Per Capita Incomes Diverged Recently?" *Federal Reserve Bank of St. Louis Review* 70 (September–October 1988), 24–30; Niles Hansen, "Economic Development and Regional Heterogeneity: A Reconsideration of Regional Policy for the United States," *Economic Development Quarterly* 2 (1988), 107–18; Orley M. Amos, Jr., "An Inquiry into the Causes of Increasing Regional Income Inequality in the United States," *Review of Regional Studies* 19 (Spring 1989), 1–12; Thomas D. Rowley, John M. Redman, and John Angle, *The Rapid Rise in State Per Capita Income Inequality in the 1980s* (Washington, D.C.: U.S. Department of Agriculture, Economic Research Service, 1991); Edward Nissan and George Carter, "Income Inequality across Regions over Time," *Growth and Change* 24 (1993), 303–19; and Rajiv Mallick, "Convergence of State Per Capita Incomes: An Examination of Its Sources," *Growth and Change* 24 (1993), 320–40.

5. Wright, *South*, 240, and Barro and Sala-i-Martin, "Convergence," table 1.

6. Coughlin and Mandelbaum, "Incomes"; Hansen, "Heterogeneity"; Lynn E. Browne, "Shifting Regional Fortunes: The Wheel Turns," *New England Economic Review* (May–June 1989), 27–46; Daniel H. Garnick, "Accounting for Regional Differences in Per Capita Personal Income Growth: An Update and Extension," *Survey of Current Business* 70 (January 1990), 29–40; Rowley, Redman, and Angle, *Inequality*; and Barro and Sala-i-Martin, "Convergence."

7. Barro and Sala-i-Martin, "Convergence."

8. The price indexes are from the 1994 *Economic Report of the President* (Washington, D.C.: Government Printing Office, 1994), tables B-59 (CPI) and B-60 (motor fuel, and fuel oil and other household fuel commodities), 335–37.

9. See Browne, "Fortunes," 36, and Rowley, Redman, and Angle, *Inequality*, 12–13.

2

EMPIRICAL STUDIES EXPLAINING INCOME

Economic literature is one place to turn to for ideas about the causes of income convergence. In this chapter, we will examine three categories of empirical literature: (1) empirical studies of income *levels*, (2) empirical studies of post-World War II income *change*, and (3) empirical studies of income divergence during the 1980s. In chapter 3 we will examine theories explaining income change. We invert the normal order of discussion, putting empirical studies ahead of theory, because the empirical literature provides a foundation for our own theory and for some of our comments on other theories.

EMPIRICAL STUDIES OF INCOME LEVELS

We look first at empirical studies of income *levels*. By learning what causes variation in income levels, we can gain insights into what causes variation in per-capita income growth rates. Two studies, one by Wheat[1] and one by Schwirian,[2] analyze variation in county median family income levels in the United States. Median family income is the alter ego of per-capita income; with minor exceptions (e.g., number of workers per family), the determinants of one are about the same as the determinants of the other.

The Wheat Study

Wheat studies variations in 1969 (1970 census) median family income among 2,706 of the nation's 3,141 counties and county equivalents. The excluded counties have abnormal characteristics (e.g., military bases, prisons) that could bias the findings.

The study begins with the observation that low-income counties tend to have

one or both of two characteristics: a southern location and ruralness. Extremely low income—low enough to qualify a county for federal designation as a distressed area—almost always requires a combination of the two characteristics. Low-income counties are concentrated in the South because of the South's legacy of slavery. Slavery gave the South an unfavorable rural-urban population mix; inefficient, labor-intensive agriculture; a relative lack of manufacturing; a high black population percentage; low educational levels; and extraordinarily low wages. Each of these characteristics hurt income.

Ruralness is the second characteristic of low-income counties. Wheat looks at the rural-urban characteristics of 198 southern counties among the 203 U.S. counties designated in 1969 as low-income counties (eligible for federal developmental assistance). In 89 percent of the low-income counties, the principal city was under 5,000 in population, whereas only 52 percent of the nondesignated southern counties had a principal city under 5,000. Only 2.5 percent of the low-income counties but 27.5 percent of the others had a city above 10,000. Only one low-income county had a city above 15,000; none had a city above 25,000. The urban population percentage was thirty percent or less in 89 percent of the low-income counties but in only 48 percent of the others. Only 1 percent of the low-income counties but 26 percent of the others were more than fifty percent urban.

The explanation is that urban counties have better occupational mixes, higher wage-salary structures (reflecting higher costs of living), and higher per-capita employment. A +.41 correlation (2,706 counties) between professional-technical-managerial employment percentage and urban population percentage confirms what we all know: cities have proportionately more professional employment than rural areas. And a +.47 correlation between the professional-technical-managerial percentage and median income verifies that professionals tend to earn more than nonprofessionals. Living costs are higher in urban areas because of higher land costs (higher housing costs), higher taxes (more facilities and services), higher business costs (higher rents, unions, higher taxes), and higher transportation costs. Per-capita employment is higher because of lower fertility (fewer children, or nonworkers) and more jobs for women and children.

A classification analysis groups the study's 2,706 counties into nine median family income intervals. For the counties in each interval, the analysis examines the average values of selected variables. The idea is to see whether, for particular variables, these average values go up and down with income. Table 2.1 summarizes the most important classification findings. The seven variables (last seven column headings) represent six factors that affect income; two variables represent the urbanization factor. Most of the labels are self-explanatory, but three need clarifying. WAGES: OPERATIVES is median annual earnings for the census "operatives" occupational category; it is a proxy for hourly wages, for which no county figures are published. Using an occupation-specific earnings figure minimizes the influence of variation in occupational mix on the variable's value. L/P (Labor Force/Population) is the labor force participation rate, a measure of the percentage of the population that earns income. POP%WIDOW-DIVOR is the population percentage of widowed and divorced women, a low-income group.

Table 2.1
Average Values of Selected Variables for Counties in Nine 1969 Income Intervals

Median Family Income Interval	Number of Counties	POP% HIGH SCHOOL GRAD	WAGES: OPERATIVES	L/P	POPULATION	POP% URBAN	POP% WIDOW-DIVOR	POP% BLACK
$11,000 up	88	61.0%	$8,123	41.5%	472,294	72.1%	5.5%	4.3%
10-10,999	141	55.1	7,440	40.6	265,329	63.9	6.0	4.5
9-9,999	250	53.9	6,988	39.9	95,476	49.1	6.2	3.2
8-8,999	496	50.5	6,514	38.8	53,332	39.2	6.5	4.0
7-7,999	549	45.4	5,907	37.6	24,894	30.1	6.7	5.7
6-6,999	514	39.9	5,219	36.5	16,051	23.5	7.1	8.8
5-5,999	417	33.0	4,583	34.8	14,488	16.4	7.5	16.1
4-4,999	192	27.6	4,078	31.8	14,779	17.9	7.5	21.4
2-3,999	59	23.7	3,444	27.3	10,544	7.1	7.1	24.2

Source: Leonard F. Wheat, *Labor Force Participation and Unemployment in American Counties* (Washington, D.C.: Economic Development Administration, 1982), Appendix A, "Determinants of Median Family Income," table A-2.

The table shows a series of perfect or near-perfect progressions: the average value (for counties in the income interval) for any given variable always or almost always increases (or else decreases) with each increase in income (bottom-to-top interval order). POP%HIGHSCHOOLGRAD (education), WAGES:OPERATIVES (wage level), and L/P (participation) show exceptionally strong relationships with income: each increase in income involves an increase in the average value (for counties in the income interval) of each of the three variables. POPULATION and POP%URBAN (urbanization) increase almost as steadily, showing just a minor aberration in the third-lowest interval. POP%WIDOW-DIVOR deviates from a steady progression—decreases this time—only for the first two income increases. POP%BLACK (race) decreases with each increase in income in all but the two highest income intervals, where increases in the black percentage are associated with large metropolitan areas (POPULATION); the positive effects of metropolitanization offset the negative effects of race.

Although the classification findings hold nothing constant, the strength of the relationships is impressive. This preliminary evidence strongly suggests that income is influenced by education, wage rates, participation (roughly equivalent to per-capita employment), urbanization, the widowed-divorced percentage, and race. And, as it happens, the observed relationships stand up very well in regression tests, which do hold other things constant. The exception is POP%BLACK. It is duplicated by POP%HIGHSCHOOLGRAD and loses significance when the latter is controlled.

The study presents several "Best-R" equations. Their variables yield the highest (second-highest in one instance) Rs that can be achieved by given numbers of significant, right-sign variables. Table 2.2 shows the best R_2, second-best R_2, best R_5, and best R_{10} equations. (The subscript says how many variables are in the equation.) Note that the ten-variable (R_{10}) equation has *two* urbanization variables (logPOPDENSITY and POP%FARM) and *three* participation variables (L/P; AGE%15-24, describing a low participation age group; and SOCIAL-SECURITY, describing retirees, who don't work). Two other variables require comment. MALE:%NEV-MARRY (population percentage of never-married males) is positive because many never-married males are young men still living with their parents; they enhance *family* income but would have the opposite effect on *per-capita* income. SOUTH-CENTRAL is a regional dummy whose negative sign may reflect, among other things, the negative effects on income of racial discrimination; the region includes the state with the highest black population percentage (Mississippi).

Four factors—education, urbanization, wage level, and labor force participation—are dominant in the equations. The consensus of the four equations seems to be that education and urbanization are the two strongest influences, with participation ranking third. This first-first-third ranking is striking because, to anticipate, our study will arrive at the same ranking.

Table 2.2
Selected Regression Equations for Predicting 1969 Median Family Income and Simple Correlations for Variables Used

		Best R_2		Second-Best R_2		Best R_5		Best R_{10}	
Variable	r	Coef.	F	Coef.	F	Coef.	F	Coef.	F
HI-SCHOOL-GRAD	+.73	+106.6	5049			+53.9	1401	+52.2	1709
logPOPDENSITY	+.46	+1270.8	1861			+984.5	1833	+845.2	1366
WAGES:OPERATIVE	+.73			+740.2	2873	+361.4	1005	+291.2	806
L/P	+.63			+192.3	1566	+117.3	1004	+109.8	1022
POP%WIDOW-DIVOR	-.33					-256.7	791	-219.8	457
POP%FARM	-.36							-24.7	508
AGE%15-24	.14							-127.3	251
SOCIAL-SECURITY	-.50							-58.9	230
MALE-%NEV-MARRY	-.08							+92.8	87
SOUTH-CENTRAL	-.32							-333.9	80
Constant		+826		-3,998		-1,128		+2108	
R		.852		.840		.933		.950	

Source: Ibid., table A-1.

The Schwirian Study

Schwirian performs a similar cross-sectional analysis of counties that substitutes 1979 (1980 census) median family income for 1969 income.[3] And he provides a second regression analysis that substitutes percentage of families below the poverty level for median income. Both analyses use all U.S. counties. Instead of a series of ever-longer best-*R* equations, Schwirian formulates the equation with the highest *R* that can be reached with an unlimited number of significant, right-sign variables. (The number of variables is 24 for the median income equation and 15 for the poverty equation.) He then provides a consolidated equation that combines all variables that measure the same factor into a compound variable—one compound variable per multivariable factor. The procedure is similar to factor analysis but uses a technique introduced by Wheat.[4]

Table 2.3 summarizes Schwirian's findings for the two consolidated equations, which use the compound variables. The median income equation has compound variables named for six factors: PARTICIPATION-RATE, RURAL-URBAN-MIX, EDUCATION, INDUSTRY-MIX, WAGE-RATE, and LOW-EARNING-POWER (population groups). The poverty equation has four of the six compound variables but omits EDUCATION and INDUSTRY-MIX. These two variables are left out because the basic equation (15 variables) has only one variable for each of these factors; there is nothing to combine into a compound variable. In place of EDUCATION the poverty equation uses the basic equation's education variable: SCHOOLING:0-8YEARS, the adult population percentage with eight or fewer years of schooling. In place of INDUSTRY-MIX the poverty equation uses LABOR%HI-MFG, the percentage of the labor force employed in high-wage manufacturing industries (all manufacturing except textiles and apparel). Because the regression coefficients for compound variables are always 1.0, the table shows no coefficients. Instead it shows the *t* values (significance) of the variables.

For the median income equation, the factors rank as follows: (1) participation rate, (2) rural-urban mix, (3) education, (4) industry mix, (5) wage rate, and (6) low-earning-power groups. The low-earnings groups combined into LOW-EARNING-POWER are military personnel, Hispanics, widowed-divorced females, and females in general. As in Wheat's study, the POP%BLACK variable is duplicated by education variables and does not get into the basic equation.

For the poverty equation, the ranking is (1) participation, (2) rural-urban mix, (3) education, (4) low-earning-power groups, (5) wage rate, and (6) industry mix. This time the low-earnings groups are females, Hispanics, and blacks. Note that both equations have the same one-two-three ranking for the three strongest factors. Also note that the three strongest factors are the same ones that were strongest in Wheat's study, although participation has moved up from third to first.

EMPIRICAL STUDIES OF POSTWAR INCOME CHANGE

The next literature category is empirical studies that examine per-capita income *changes* since World War II. Some of these studies also cover earlier periods or

Table 2.3
Values of *t* from Consolidated Equations for Predicting 1979 Median Family Income and Percentage of Families Below Poverty Level

	Dependent Variable	
Variable	**Median Income**	**Poverty Percentage**
PARTICIPATION-RATE	56.4	57.2
RURAL-URBAN-MIX	55.0	40.7
EDUCATION	36.3	
INDUSTRY-MIX	24.9	
WAGE-RATE	12.9	12.0
LOW-EARNING-POWER (groups)	7.6	22.7
SCHOOLING:0-8YEARS		31.5
LABOR%HI-MFG		5.6
Adjusted R^2	.82	.83

Source: Kent P. Schwirian, *Determinants of County and City Median Family Income Levels and Poverty Rates* (Washington, D.C.: Economic Development Administration, 1991), tables 3.11, 3.12.

deal mainly with topics other than income; they nevertheless have valuable material. The significant postwar studies are those of Hanna, Perloff et al., Browne, Garnick and Friedenberg, Chinitz, Persky, Wright, and Lemann.

The Hanna Study

Hanna's *State Income Differentials, 1919–1954* studies changes in per-capita income.[5] The work is largely descriptive, but it does examine the relationship of certain factors to income. Hanna finds that the southern states have the most unfavorable occupational compositions and industry compositions. Hanna's industry analysis is limited to manufacturers, wholesale trade, and government, however, so the all-important effect of agriculture as a component of industry mix gets short shrift. Along with the mountain states, the southern states also tended to have the largest percentages of their populations in non-working-age (non-income-earning) age brackets.

Looking at cities, Hanna finds that the average median incomes for cities in different size classes tend to increase as the size increases. He doesn't say so, but the implication is that the South—the most rural region—has an unfavorable city size hierarchy. A related finding is that manufacturing employment as a percentage of total employment increases with city size. This finding supports conventional wisdom, which recognizes the contribution of manufacturing growth

to population growth. (To anticipate, our study finds that manufacturing change is only a weak determinant of income change. We think part of the reason is that manufacturing change is largely duplicated by a broader and more powerful influence, urban growth, which we control.)

Hanna stops short of saying that occupational mix, industry mix, age mix, and city size mix (or urbanization) are causes of per-capita income variation, but he certainly implies this. It follows that a study of the causes of income change should include these factors or surrogates for them. (Urban population percentage is a surrogate for occupational mix and city size mix.)

The Perloff Study

The Perloff study, named for its principal author, is really the work of Perloff, Dunn, Lampard, and Muth.[6] It looks at 1920–55 income change. Although most of the analysis is descriptive, the authors do offer conclusions about things that affect income levels and that cause changes in those levels. The conclusions are based on simple correlations and on income comparisons of different population groups. Per-capita income, they state, has two main determinants; the second has several subordinate determinants. The two main determinants are (1) per-capita employment and (2) average earnings of the employed persons. Regarding the first, high-income states tend to have proportionately many working-age persons and proportionately high percentages of the working-age persons in the labor force.

The average earnings of employed persons depend heavily on industry mix and racial mix. Where industry mix is concerned, the most important industry is agriculture. The states with the lowest per-capita incomes tend to be those with the highest percentages of people working on farms. And differences among the states in their rates of agricultural contraction were the strongest force behind the 1920–55 narrowing of state income differences: the agricultural states contracted more. A small but significant correlation exists between median income and the percentage of the labor force employed in manufacturing. (Something Perloff et al. do not note is that the correlation is weak because income is high in the West, where manufacturing is weak.) A stronger correlation exists between income and the percentage of the labor force employed in business service industries. (This percentage would seem to reflect the urban population percentage and occupational mix.)

Race is "extremely important."[7] In 1949 the median income level of blacks living on farms in the Southeast was only about half that of whites living on farms ($485 compared to $933). Note that this comparison controls the effect of industry mix. Nonfarm sectors also had large black-white income differences, though not as large as the farm differences. And the farm and nonfarm sectors in other regions likewise showed black-white differences. For example, the median income for blacks in the Mid-Atlantic region was $1,344, compared to $2,330 for whites.

The Browne Study

Browne provides the most thorough analysis of postwar income changes.[8] She investigates the causes of per-capita income convergence between the late 1950s and 1980. Convergence during this period, she observes, is an extension of an uneven trend going back to the turn of the century. To explain convergence, she identifies three primary sources of income differences: (1) differences in per-capita employment, or the percentage of the population that is employed, (2) differences in wage-salary earnings per employed person, and (3) differences in transfer payment and property income per person (including those not employed). In essence, this approach decomposes income into the following elements:

1. Wage-salary income
   a. Nonearners: population percentage not employed (zero earnings).
   b. Earners: wage-salary earnings per worker.
2. Transfer and property income per person (employed and unemployed).

Specific influences, such as farm population decline and wage level increases, are subsumed under these headings. Browne analyzes change at the regional level, using the nine census divisions as her regions.

According to Browne, the most important cause of income convergence was convergence in wage-salary earnings per worker (1*b* above). This convergence resulted partly from farm population percentage declines and partly from changes in nonfarm wage levels. In 1967, income per worker in farming was only 52 percent of that for the nonfarm sectors. Shifts from farm to nonfarm employment therefore raised per-capita income. During the first decade or so of the study period, farm population declines caused earnings per worker to rise considerably in the South and somewhat less in the other comparatively agricultural regions—the Plains and the Rocky Mountain states. Farm-to-urban shift effects were dominant in the 1960s but weakened in the 1970s because of agriculture's diminished size.

Around the late 1960s, southern gains in nonfarm earnings displaced farm-to-urban population shifts as the more important cause of income convergence. Ironically, the nonfarm earnings gains resulted from the decline in agriculture. The bigger the farm sector, the smaller the growth in nonfarm wages. In the 1950s and 1960s, when the farm sector was still relatively big, the flow of labor from farms to cities increased the nonfarm labor supply and thereby restrained wages. But as the farm labor pool dried up, labor shortages developed and nonfarm wages climbed. Declines in southern unemployment rates also contributed slightly to the wage gains. Browne does not say so, but the unemployment declines would seem to be a facet of the decline in farm-to-urban migration, the source of surplus labor.

A striking feature of the nonfarm earnings picture was a decline in western earnings relative to the national average. In 1958, the Pacific subregion of the West had the highest nonfarm earnings per worker: 114 percent of the national

average. Next highest were two Manufacturing Belt subregions: East North Central (108 percent) and Mid-Atlantic (107 percent). By 1977, the Pacific states had fallen 5 percentage points to 109 percent. In contrast, the Mid-Atlantic states went up 2 points to 109 percent and the East North Central states fell just 2 points to 106 percent. Browne's findings support a theory that we will advance: that cost of living declines in the West led to wage-salary declines that lowered relative per-capita income.

The second of the three income elements Browne analyzes is per-capita employment. She finds that rising employment/population ratios in the low-income regions contributed importantly to convergence. In 1960 per-capita employment was below the national average in five regions. These were the five most agricultural regions, which were also the lowest income regions (South Atlantic, East South Central, West South Central, Plains, and Mountain). These same five regions, and only these five, experienced above-average 1960–76 gains in per-capita employment.

The gains in per-capita employment came from two main sources. The more important source was increases in the labor force participation rate (i.e., the percentage of the population that is in the labor force, where labor force is the sum of employed and unemployed workers). Browne doesn't say so, but participation rates are higher in cities (jobs for women and teenagers) than in farm areas; rising participation is a predictable outcome of farm-to-urban population shifts. Alluding to the civil rights movement, Browne credits reduced employment discrimination and increased access to education with some of the participation rate gains. The second main source of per-capita employment gains was increases in the ratio of working-age population to total population. Only in the 1960s did this source have much effect. The primary reason for working-age population percentage increases was birthrate declines; these declines were greatest in the southern and mountain states. Though Browne again doesn't say so, fertility rates tend to be higher for farm families; a farm-to-urban population shift tends to reduce the ratio of children to population and thereby to increase the working-age percentage.

The third of Browne's three income elements is transfer payments and property income (dividends, interest, and rent). This element is heavily influenced by the earnings-per-worker element. In general, states with high earnings per worker tend to have high transfer and property income per person. For property income, the connection is obvious: people with high earnings have more money to invest, so they earn more dividends, interest, and rent. For transfer payments, the connection with earnings is less direct. Government retirement programs base payments on past income, so states with high earnings show high payments to retirees. Moreover, states with high earnings per worker have high tax bases and can thus afford more generous public assistance and unemployment compensation programs. The main effect of changes in transfer and property income has been to magnify the effects of changes in earned (wage-salary) income.

The Garnick-Friedenberg Study

Garnick and Friedenberg look at per-capita income changes for 1929–79; they emphasize 1949–79.[9] Like Browne, the authors aggregate at the regional level, but they use the eight-region Bureau of Economic Analysis configuration instead of the Census Bureau configuration. Also like Browne, they decompose income into elements, as follows:

1. Labor and Proprietors' Income (Browne's wage-salary income)
   a. Changes due to industry mix changes.
   b. Changes due to employment/population ratio changes.
      (1) Changes in employed percentage of working-age population.
      (2) Changes in working-age percentage of total population.
   c. Changes in industry-mix-adjusted regional wage rate.
2. Other Income
   a. Changes in per-capita transfer payment income.
   b. Changes in per-capita property income (dividends, interest, rent).

The analysis of "other income" is not too informative, since it uses income changes to explain income changes. But the rest of the analysis offers important insights. The three facets of labor and proprietors' income jointly explain about three-quarters of the 1940–79 narrowing in state per-capita income. About half of the overall convergence (and two-thirds of the convergence in labor and proprietors' income) resulted from 1979's more uniform industry mix; industry includes agriculture. Another one-tenth of the overall convergence came from more uniform ratios of employment to working-age population, one-twentieth came from more uniform working-age population percentages, and one-tenth came from more uniform regional wage rates.

Browne, you will recall, concluded that most change in transfer and property income depends on change in labor and proprietors' income. Applying this conclusion to the Garnick-Friedenberg findings, we can estimate that five-eighths of the overall convergence (or one-half plus half of labor and proprietors' income's one-fourth) resulted from changes in industry mix. The authors attribute most of this industry mix effect to employment shifts from farming to higher-paying industries. But in New England in the 1950s, the effect of the textile industry's shift from New England to the South was visible: New England acquired proportionately more durable goods manufacturing (high wages) and proportionately less nondurable goods manufacturing (low wages).

The Chinitz Analysis

Chinitz's analysis contains an interesting interpretation of per-capita income growth in the South.[10] This interpretation is notable for the emphasis it puts on

industrialization as the main source of per-capita income growth in the South during this century. Chinitz observes that southern income has converged on the national average. In his opinion, the most important factor behind income growth in the South was manufacturing growth. And the most important factor behind the manufacturing growth was the South's low wages. Wages were low because labor was being released from agriculture. (Here Chinitz overlooks the fact that southern wages were always low, because of the effect of slavery; surplus labor released by the technological revolution in agriculture merely helped *keep* southern wages low. But this is a quibble.) Union weakness also contributed to the South's low wages. And beyond wages, the South offered low business taxes and other investment incentives.

Chinitz also cites a more abstract cause of convergence, a cause that goes beyond the South to take in all regions. This abstract cause is increased regional homogeneity: the regions became more alike in their socioeconomic makeup. Chinitz specifically mentions the more even geographic distributions of manufacturing, agriculture, black population, urban and rural population, and poverty. The homogeneity thesis has great merit, as far as it goes; in fact, it resembles our own explanation of convergence, both in the abstract and in its few particulars.

But the homogeneity explanation does have three weaknesses. First, the list of particulars omits two of the most important aspects of the observed increase in homogeneity—education and per-capita employment. To anticipate, we shall see that these are far more important than manufacturing, the aspect that Chinitz considers most important. Second, the references to agriculture, race, and poverty are too superficial. They ignore the underlying causes of the South's historically low income—slavery and its replacement, the sharecropper-tenant system of agriculture ("sharecropping"). The thesis thus overlooks the profound contribution of sharecropping's collapse to the increases in homogeneity. Third, the homogeneity explanation partly conflicts with the reality of slow income growth in the West. The West gained on all other regions in per-capita manufacturing, and it gained on the Manufacturing Belt in urbanization. Also, the West absorbed proportionately fewer blacksg (poverty persons) than the Manufacturing Belt. Why, then, was the West's 1950–87 income growth rate lower than those of the Manufacturing Belt (which had the slowest growth and the highest black in-migration) and every other region?

The Persky Study

In sharp contrast to Chinitz, Persky puts the focus on agriculture. He provides a detailed picture of southern off-the-farm migration.[11] His article is useful both for its detailed description of developments in southern agriculture and for its assessment of the per-capita income effects of these developments.

Persky observes at the outset: "There is a general consensus among regional economists that the transfer of labor between the farm and non-farm sectors of the South has been the major source of the region's growth in per-capita income and its ability to 'catch up' with the rest of the nation."[12] He is wrong about the

"general consensus" but is otherwise on track. The income change literature, except for Perloff's analysis, tends to downplay and often ignores southern farm-to-urban migration. And it almost totally overlooks the main cause of that migration—sharecropping's collapse. The studies that mention sharecropping deal with other topics, such as agriculture, migration, race, poverty, and the southern economy. Persky's study, for example, deals with migration. Still, Persky's point about the income effects of farm migration is well taken.

Persky observes that southern agriculture between the Civil War and World War I was a mixture of three farm types. At one extreme were small owner-operated farms that produced modest amounts of cash crops like cotton and tobacco along with substantial quantities of food crops; the owners were predominantly white. At the other extreme were large plantations based on some version of the sharecropper-tenant farmer system. The tenants and sharecroppers were mainly black, and they were subject to considerable supervision. Between these extremes were substantial numbers of small tenant farms whose operators were not directly supervised. The operator-tenants were usually white but sometimes black. Mounting debt among small owners led to an upswing in tenancy in the 1920s. By 1930, tenants constituted 55.5 percent of the *operators*. The combined total of tenant-operators and plantation tenants (including sharecroppers) came to around five-eighths of all southern farmers.

The Great Depression was the beginning of the end of tenant farming; it was also a period of contraction for overall farming in the South. At first, however, the number of tenants and farmers increased: a return-to-the-farm movement among unemployed workers carried into 1935. Two developments then caused the tide to reverse. First, New Deal acreage retirement and crop control programs reduced the demand for farm labor. At the same time, provisions for sharing payments with tenants gave landlords an incentive to begin replacing tenants with hourly wage labor. Second, the tractor, which arrived in the 1920s, began to proliferate. Flatter lands made the western part of the South, particularly Oklahoma and Texas, better suited for tractor cultivation; hence, most of mechanization's effect was in the Southwest. Tractors reduced the need for labor and encouraged farm consolidation: consolidation spread the tractor investment over more acreage, making the tractor affordable by reducing its per-acre cost.

In the 1940s the off-the-farm movement accelerated. World War II pulled many workers from agriculture into the armed forces or defense industry jobs. The tractor continued to multiply and moved eastward, displacing still more workers. And the mechanical cotton picker, though still in its early stages of introduction, began to make itself known. Where used, it nearly eliminated the need for unskilled labor. The shift towards hourly wage workers, corporate farming, and mechanization was now well under way.

This shift intensified in the 1950s. The mechanical cotton picker came into widespread use. In the Southeast, the terrain was too hilly for intensive mechanization; farmers began shifting from cotton to more suitable crops. Food crops, livestock, and poultry became much more important. The new products, like cotton under mechanization, required less labor and favored larger owner-operated farms. The tenants were forced out.

All the foregoing developments involved "push" influences: things that pushed

the tenants and small operators off the farms. But off-the-farm migration also reflected "pull" influences: higher wages in urban occupations. Persky does not specifically mention southern industrial growth as a factor that magnified the pull influence, but the implication is there.

The Wright Study

Gavin Wright's book, *Old South, New South: Revolutions in the Southern Economy Since the Civil War*, is the definitive study of economic change in the South since the Civil War.[13] This study enlarges the picture given by Persky and provides more depth and detail. The plantation system of agriculture did not wither away when slavery was abolished, but the landowners did have to develop a new economic relationship with the former slaves. Early attempts to establish work gangs of paid laborers failed: the freedmen resisted this neoslavery arrangement by moving from place to place in search of better conditions. Work "squads" based on varying degrees of kinship followed, but worker resistance to discipline led to increasing worker autonomy. By 1880 most planters agreed that the best arrangement was one that let each family work by itself in separate fields or on a separate farm. This new arrangement entailed not only subdividing the fields but decentralizing the housing of the freedmen. The former slave quarters near the master's house were replaced by widely scattered cabins, each located on its own subdivision.

Meanwhile, the planters lacked the means to pay cash wages, despite the early attempts. A rapid decline in the price of cotton after 1865 caused cash wage agreements to quickly fall out of favor. Sharecropping arose as a more practical arrangement. In return for their labor, the workers received shares of their crops. The transition to sharecropping was facilitated by the fact that sharecropping agreements between planters and white tenants had existed before the Civil War.

As things evolved, there was not always a clear distinction between tenants and sharecroppers; the two terms tended to be used interchangeably. (Sharecropping is actually a form of tenant farming: the *American Heritage Dictionary* defines *sharecropper* as "a tenant farmer who gives a share of his crop to the landlord in lieu of rent.") Cash tenants existed before the Civil War and grew in number after the war. But share tenancy was more common. It, in turn, had many intermediate forms between autonomous (unsupervised) tenancy and sharecropper arrangements that involved some degree of supervision by the planter. One intermediate form of share tenancy was reached when the tenant provided his own implements and his own mule. This step toward independence entitled the tenant to a larger share of the crop. But whatever the precise arrangement—here we interpolate—the essential fact was that the sharecroppers and tenants were abysmally poor. Since the South remained predominantly agricultural after the Civil War, this sharecropper poverty was the basis for southern poverty.

In large parts of the South, small owner-occupied farms rather than plantations and tenant farms were the rule; this type of farming antedated the Civil War. The proprietorship farms tended to be marginal, generally not much more lucrative than tenant farms. The farmers were usually white, although the number of black

proprietorships grew steadily. Whites, for their part, made inroads into tenant farming. Two-thirds of the South's share tenants were white by 1900. The poverty of southern agriculture thus extended to whites as well as blacks.

The sharecropper-tenant system weakened in the 1930s. New Deal "soil conservation" programs, designed to bolster farm prices by taking land out of production, provided land withdrawal payments to farmers. These payments had to be shared by the landlords and tenants, provided that tenant-sharecropper labor was used. To avoid payment sharing, some landlords switched from tenants to hourly wage labor. Land withdrawal forced additional tenants off their farms. At the same time, tractors slowly came into use. As southern agriculture became more mechanized, less labor intensive, still more tenants and sharecroppers were displaced.

Even so, sharecropping remained fairly strong until World War II. Abundant cheap labor—blacks and poor whites—was a barrier to mechanization. But during World War II the South's farm population fell 22 percent; young men were inducted into the military or took factory jobs in the North and South. After the war, the farm outflow continued, if at a slower rate. Agricultural labor shortages became acute. The labor shortages gave International Harvester the incentive to develop a commercially successful mechanical cotton picker.

Mechanical cotton pickers were sold as early as 1941, but various problems kept sales to insignificant levels before 1950. By the late 1940s, however, "enough cotton planters . . . were potentially in the market for a mechanical harvester that it was worth the investment of time and money to satisfy their demand."[14] Soon John Deere and Allis Chalmers joined International Harvester in the market. Machine-harvested cotton went from 5 percent of the crop in 1950 to 50 percent in 1960 to over 90 percent before 1970. Sharecroppers, tenant farmers, and hired laborers of both races departed from southern farms in droves. Many went North, others to southern cities. Mechanized tobacco harvesting followed in the 1970s, extending the off-the-farm movement.

Wright draws no income conclusions from these developments, but such conclusions are easily drawn. Those who left the farms were (*a*) rural residents, (*b*) employed in agriculture, (*c*) poorly educated, and (*d*) proportionately more black than the nonsouthern population. The migration of sharecroppers to southern cities and nonsouthern regions therefore had profound effects on rural-urban mix, industry mix, educational levels, and racial mix. (The last two characteristics were affected only by North-South migration.) These changes were bound to affect per-capita income in the South.

The Lemann Study

In *The Promised Land* Lemann more or less picks up where Wright leaves off.[15] This book describes "the Great Black Migration" that resulted from the technological revolution in southern agriculture. Five million blacks left the South's farms and migrated to the North between 1940 and 1970. Whereas 77 percent of all black Americans lived in the South in 1940, only half of America's

blacks were southern in 1970. The biggest single factor behind the great migration was the mechanical cotton picker. "Within a few years after the end of World War II, the mechanical picker was coming into general use on the plantations, and the sharecropper system was ending."[16] By the 1960s, bigger and better farm equipment, including eight-row tractors and four-row pickers, were being introduced. Abundant farm labor was no longer needed.

Chemical agriculture, particularly chemical weeding, helped the new equipment drive the sharecroppers off the farms. In the mid-1950s, chemical companies began trying to market powerful new defoliants for weeding the cotton fields. The planters were at first reluctant to use them, but gradually the new chemicals gained acceptance. Herbicides ultimately made hand "chopping" (hoeing) of weeds unnecessary. The first herbicide to come into widespread use was Du Pont's Treflan, a pre-emergent weed killer. It was followed by postemergent weed killers, including one that attacked Johnson grass, which choked the cotton in late summer. Other chemicals made cotton plants grow bigger; still others loosened the cotton bolls just in time for picking.

A 1967 development provided the coup de grace for the remnants of the sharecroppers. At this time sharecropping was already moribund, but many former sharecroppers were still working as day laborers, chopping weeds. Then the federal government brought farm labor under the minimum wage. Southern planters suddenly found they would have to pay $1.15 an hour to workers they had been paying $3.00 per day. The planters reacted quickly: they switched to chemical defoliants. The remaining sharecroppers were out of work. For all practical purposes, sharecropping and other labor-intensive agriculture was dead. And—here we again interpolate—so was the poverty group that for so long had depressed southern income.

Out of work, the sharecroppers had to go somewhere. Many moved to cities in the South, but southern cities were not the most inviting: racial prejudice and discrimination ran high. The sharecroppers knew people, often relatives, who had moved to Chicago and other northern cities. These people painted a rosy picture:

> In Chicago you could make [good wages] working in a laundry, or a factory, or a restaurant or a hotel, or one of the big mail-order houses like Spiegel and Montgomery Ward, or, if you were a man, in the stockyards. You could get overtime. Some of these jobs were supposed to be as hard as picking cotton, but people were making sums unheard of among black unskilled workers in Mississippi.[17]

And so the Great Black Migration to the North proceeded. It was one of history's largest and quickest mass internal population movements. Neither the Irish nor the Italians nor the Jews nor the Poles had migrated to America in such large numbers. In just the last decade of the migration, the 1960s, Chicago's black population grew by 300,000. Soon *ghetto* became a household word. Race and poverty had moved North.

EMPIRICAL STUDIES OF THE DIVERGENCE OF THE 1980S

We saw in chapter 1 that convergence temporarily became divergence between 1978 and 1988, after which income convergence resumed: each year after 1988, through 1993, had less income inequality than the preceding year. Five studies examine the divergence of the 1980s and provide reasons for it. The studies are by Coughlin and Mandelbaum; Hansen; Browne; Garnick; and Rowley, Redman, and Angle. (The last four studies include Alaska, Hawaii, and the District of Columbia, whereas the Coughlin-Mandelbaum study and our own cover only the 48 contiguous states. These coverage differences explain why the Hansen, Browne, Garnick, and Rowley-Redman-Angle studies identify 1979, not 1978, as the reversal year. Alaska, Hawaii, and the District all converged on the national average in 1979.) Several adverse subtrends identified by the studies ended in 1988 or 1990; this is why convergence resumed.

The Coughlin-Mandelbaum Study

Coughlin and Mandelbaum focus on income developments in the final decade of our study period.[18] They show that the long-run trend toward greater income equality among states reversed after 1978. Their goal is to determine why inequality increased during 1978-87. They begin by grouping the 48 contiguous states into 10 upwardly divergent ones, 10 downwardly divergent ones, 4 upwardly convergent ones, 6 downwardly convergent ones, and 18 that had no substantial change in per-capita income. The analysis emphasizes the first two categories: states that *diverged*, either upward from the U.S. average (by increasing in relative income) or downward (by decreasing in relative income).

All states with comparatively fast income growth—the 10 upwardly divergent ones and the 4 upwardly convergent ones—were on the East Coast (if Vermont is counted as an East Coast state). A resurgence of manufacturing growth, particularly in computer-related and defense industries, caused the income growth in five New England states. Rhode Island, New York, and New Jersey did not share in the manufacturing resurgence; their gains were in construction and services. The reasons for this construction-service growth are unclear. The authors think, however, than an increase in federal grants-in-aid may have helped. The southern coastal states grew through continuation of their long-run upward trend in industrialization.

But why didn't this upward manufacturing trend bring similar 1978–87 income gains to the other southern states and to the West? The authors think the answer lies not in positive factors affecting the coastal states but in negative factors affecting the interior states. Two such factors are the energy price decline of the 1980s and the farm crisis of the 1980s.

Beginning with the 1973 energy crisis (OPEC oil embargo) and continuing beyond the 1979 energy crisis (Iranian revolution), energy prices climbed rapidly;

they peaked in 1981. Energy prices then fell steadily through 1987. Most energy-rich states had below-average incomes in 1978, so the decline in energy prices tended to cause income divergence. Coughlin and Mandelbaum identify eleven states where earnings from oil and gas extraction and coal mining accounted for at least 3 percent of total earnings for 1981, the peak energy price year. Downward divergence in eight of these energy states explains the divergence in 8 of the 10 downwardly divergent states. The 8 downwardly divergent energy states are West Virginia, Louisiana, Oklahoma, Texas, New Mexico, North Dakota, Montana, and Utah. Income also fell in a ninth energy state, Wyoming, but it began with above-average income; convergence resulted.

The other two downwardly divergent states were Indiana and Idaho. These two states seem to have been affected by the second interior factor, the farm crisis. Indiana was probably hurt by sickness in the steel industry too, although the authors do not mention this.

The second interior factor, again, was the farm crisis of the 1980s. Farm exports grew rapidly in the 1970s. But after peaking in 1981, they fell sharply in 1982 and thereafter. This decline affected not only farmers but firms linked to agriculture: suppliers of farm equipment and fertilizer, firms that transport farm products, and food processors. (It isn't clear why food processors would be affected, unless the authors are suggesting that exports of processed food fell too.) Twelve states derived at least 4 percent of their earnings from agriculture during the peak export year, 1981. Four of these states (Iowa, North Dakota, Montana, and Idaho) were downwardly divergent, one (Vermont) was upwardly divergent (because of manufacturing growth), and seven had no substantial change. Two of the downwardly divergent farm states (North Dakota and Montana) were among the 8 downwardly divergent energy states. The 2 new downwardly divergent states (Iowa and Idaho) account for the last 2 of the 10 downwardly divergent states.

Other Divergence Studies

The other four studies of income divergence during the 1980s have conclusions that more or less agree with those of Coughlin and Mandelbaum. Hansen, for example, attributes the divergence to the Northeast economic boom, the energy price slump, and the agricultural depression.[19] The Northeast's boom, he concludes, resulted from increases in defense spending and, in New England, growth in high-tech industry. The biggest relative losses in income were in the Rocky Mountains and Southwest regions (BEA configuration). Energy price declines were responsible. In the Plains, declines in relative income resulted from hard times in agriculture.

The studies by Browne[20] and Garnick[21] both use the income disaggregation methods these authors used in their earlier studies. Browne finds that changes in per-capita employment contributed less to the income divergence than did changes in earnings per worker. But New England and the Mid-Atlantic states show relative gains in per-capita employment. She attributes these gains to rising

defense spending, the beneficial effects of oil price declines in a high-oil-consumption region, growth in high-tech industries, and growth in national and international financial services. These influences also boosted Northeast earnings per worker. In the West South Central and Mountain regions, the oil price declines adversely affected per-capita employment and earnings per worker in several industries, including construction, real estate, and financial services.

Garnick concludes that the divergence is fully explained by three components of labor and proprietor's income: (1) regional earnings per worker, (2) per-capita employment, and (3) industry mix. Earnings per worker changes caused about five-tenths of the joint effect, per-capita employment changes about four-tenths, and industry mix changes about one-tenth. Garnick has relatively little to say about the underlying causes of the changes in these income components. But he does refer to falling unemployment rates in New England and the Mideast and to the reversal in the 1980s of previous unsustainable changes that had helped oil, gas, and related industries. The oil-gas reversal hurt the Southwest and Rocky Mountain regions.

Rowley, Redman, and Angle look at five potential sources of 1979-89 income divergence: manufacturing dependence, the farm crisis, the energy price crisis, high-technology manufacturing, and producer services.[22] For each of these five sources, the authors identify 6 to 12 states that rank highest in the relevant characteristic. Then, looking at the characteristics one at a time, the authors calculate the 1979–89 increases in the coefficient of dispersion with and without a particular group of highest-ranking states included in the calculation. The difference in the two increases measures the contribution of a particular group of states (i.e., a particular cause). The authors find that six high-tech manufacturing states and ten producer services states contributed the most to divergence. All of the high-tech states and 6 of the 10 producer services states were in the Northeast; these states were associated with the Northeast's short-lived economic boom of the 1980s.

Divergence as a Temporary Phenomenon

The divergence examined in the five studies of recent income change has ended. We view it as an aberration, not as a lasting reversal of the long-run trend towards convergence. All five of the short-run studies reviewed here attribute significant divergence effects to the energy and farm crises, wherein 1981 peaks were the starting points for downward trends. These crises are over; further divergence resulting from these influences cannot be anticipated. The consumer price index (CPI) for fuel oil and other household fuel peaked at 104.6 in 1981, bottomed out at 77.6 in 1986, and then moved to 90.3 in 1993.[23] The CPI for motor fuel followed the same pattern: it peaked at 108.5 in 1981, bottomed out at 77.1 in 1986, and moved to 98.0 in 1993.[24] The producer price index for energy displayed similar behavior: it peaked at 101.5 in 1981, bottomed out at 59.8 in 1988, then rose to 78.0 in 1993.[25]

In agriculture, exports peaked in 1981, but the producer price index for farm products did not peak until 1984. The index went from 105.5 in 1984 to 92.9 in 1986 to 107.0 in 1993.[26] Farm price cycles will surely recur, but the point is that the farm crisis of the 1980s is over.

The Rowley-Redman-Angle study concludes that growth in high-tech manufacturing and producer services contributed more to divergence than did the energy and farm crises. Regardless of whether defense contracts or intrinsic qualities of the industries caused them to grow, the expansion was concentrated in the Northeast, particularly New England. But here again the divergence effect has run its course. A recent study by Fieser shows that the regional unemployment rate for the Northeast (census region) fell from a high of 8.7 percent in 1983 to a low of 4.1 percent in 1988. The Northeast's 1988 unemployment was well below the rates of the other three census regions, which averaged 5.9 percent. At this point the income divergence subtrend (the Northeast subtrend) ended. By 1990 the Northeast's unemployment rate was back to the national rate; by 1991 the Northeast's unemployment rate was 7.6 percent, appreciably higher than the rates of the other three regions (6.7 to 7.0 percent); and by July 1992 the Northeast's unemployment rate was 8.5 percent, compared to rates of 6.8 to 8.0 percent in the other three regions.[27] In other words, the Northeast went from a 1988 unemployment rate that was far lower than those of all other regions to a 1992 rate that was appreciably higher than those of all other regions.

The Northeast's two divisions—New England and Middle Atlantic—were the two highest of the nine census divisions in 1987 per-capita income. The Northeast's exceptional economic boom of the mid-1980s pushed these two high-income regions even higher, so income divergence resulted. But the 1988–92 reversal of the 1983–88 Northeast unemployment rate trend represents a resumption of the long-run trend toward income convergence: slower income growth in the high-income regions.

Some additional evidence from Fieser's study emphasizes this point. Fieser presents a U.S. map comparing 1989–92 industry sector unemployment rate changes for the nine census divisions. Using dark-to-light shading to depict five unemployment rate intervals, the map shows for each division the unemployment rate increase of the industrial sector having the biggest increase. The New England and Middle Atlantic divisions (transportation and manufacturing, respectively) have the two highest sectoral unemployment rate increases (6.0 to 6.9 percentage points). Two of the four divisions with the lowest 1987 per-capita income—East South Central and Mountain—have the lowest sectoral (manufacturing and finance-insurance-real estate) unemployment rate increases (2.0 to 2.9 percent).[28] A resumption of income convergence is again implied.

Since the trend reversal came in 1988, two years before the business cycle downturn of 1990, these developments cannot be written off as cyclical effects. Rather, the divergence of the 1980s was a short-run deviation from long-run trend, like the similar deviation of the 1920s. It bears repeating that the coefficient of variation (CV) declined steadily from 0.173 in 1988 to 0.142 in 1993 (table 1.2).

NOTES

1. Leonard F. Wheat, *Labor Force Participation and Unemployment in American Counties* (Washington, D.C.: Economic Development Administration, 1982), Appendix A, "Determinants of Median Family Income."

2. Kent P. Schwirian, *Determinants of County and City Median Family Income and Poverty Rates* (Washington, D.C.: Economic Development Administration, 1991).

3. Ibid.

4. Leonard F. Wheat, "The Determinants of 1963–77 Regional Manufacturing Growth: Why the South and West Grow," *Journal of Regional Science* 26 (1986), 643–44.

5. Frank A. Hanna, *State Income Differentials, 1919–1954* (Durham: Duke University Press, 1959).

6. Harvey S. Perloff, Edgar S. Dunn, Jr., Eric E. Lampard, and Richard F. Muth, *Regions, Resources, and Economic Growth* (Baltimore: Johns Hopkins Press, 1960).

7. Ibid., 601.

8. Lynn E. Browne, "Narrowing Regional Income Differentials," *New England Economic Review* (September–October 1980), 35–54, and (November–December 1980), 40–57.

9. Daniel H. Garnick and Howard L. Friedenberg, "Accounting for Regional Differences in Per Capita Personal Income Growth, 1929–79," *Survey of Current Business* 62, no. 9 (1982), 24–34.

10. Benjamin Chinitz, "The Regional Transformation of the American Economy," *Urban Studies* 23 (1986), 377–85.

11. Joseph Persky, "Push and Pull in Migration from Southern Farms," *Review of Regional Studies* 2, no. 2 (1972), 49–62.

12. Ibid., 49.

13. Gavin Wright, *Old South, New South: Revolutions in the Southern Economy Since the Civil War* (New York: Basic Books, 1986).

14. Ibid., 243.

15. Nicholas Lemann, *The Promised Land: The Great Black Migration and How It Changed America* (New York: Knopf, 1991).

16. Ibid., 50.

17. Ibid., 8.

18. Cletus C. Coughlin and Thomas B. Mandelbaum, "Why Have State Per Capita Incomes Diverged Recently?" *Federal Reserve Bank of St. Louis Review* 70, no. 5 (1988), 24–36.

19. Niles Hansen, "Economic Development and Regional Heterogeneity: A Reconsideration of Regional Policy for the United States," *Economic Development Quarterly* 2 (1988), 107–117.

20. Lynn E. Browne, "Shifting Regional Fortunes: The Wheel Turns," *New England Economic Review* (May–June 1989), 27–40.

21. Daniel H. Garnick, "Accounting for Regional Differences in Per Capita Personal Income Growth: An Update and Extension," *Survey of Current Business* 70 (January 1990), 29–40.

22. Thomas D. Rowley, John M. Redman, and John Angle, *The Rapid Rise in State Per Capita Income Inequality in the 1980s* (Washington, D.C.: U.S. Department of Agriculture, Economic Research Service, 1991).

23. *Economic Report of the President, 1994* (Washington. D.C.: Council of Economic Advisers, 1994), table B-60.

24. Ibid.

25. Ibid., table B-65.

26. Ibid., table B-66.

27. John B. Fieser, "The Structure of the 1990–1991 U.S. Recession" (paper prepared for 1993 Annual Meeting of the Association of American Geographers, Atlanta, April 7, 1993), figures 3, 11.

28. Ibid., figure 18.

3

THEORETICAL EXPLANATIONS OF INCOME CONVERGENCE

The literature reviewed in chapter 2 is essentially empirical, notwithstanding some elements of theory, inference, and speculation. But the literature of economics also includes *theory* that can be applied to explaining why per-capita incomes are converging. In particular, the literature includes (1) traditional neoclassical factor mobility and allocation theory, as epitomized in Heckscher-Ohlin trade theory, and (2) Mancur Olson's institutional sclerosis theory. And beyond these theories, we offer (3) our own empirically based theory. In our case, "theory" might be a misnomer: our explanation of convergence is quasi-historical, relying mainly on knowledge of American social and economic history rather than on abstract ideas. Yet because our interpretation of economic history is an alternative to neoclassical and sclerosis theory, it seems reasonable to call it theory.

NEOCLASSICAL AND HECKSCHER-OHLIN THEORY

The essence of neoclassical allocation and factor mobility theory is that things go where they can get the best deal. The movement of production factors (labor and capital) and trade (goods and services) to the places offering the best price or rate of return tends to equalize prices (including the price of labor) and rates of return. Under neoclassical theory, as amplified and expanded in the trade theory of Heckscher and Ohlin, workers, capital (production), and commodities flow between countries or regions in response to regional differences in wages, production costs (or rates of return on investmeni), and commodity prices.[1] The movement continues until the resulting changes in supply (of labor and capital) and demand (for goods) equalize prices and rates of return. Perfect equalization should not be expected, because regional differences in climate, natural resources, and other things permit some price differences to remain. But the tendency toward equalization—price convergence—will be strong.

Applied to state and regional per-capita income, this theory implies that (1) workers should migrate from the South to regions with higher wages, creating labor shortages in the South and thereby driving up wages, (2) manufacturing, attracted by low labor costs, should flow into the South, generating new income, and (3) goods, manufactured more cheaply in the South, should flow from the South to other regions, bringing still more income to the South. Regions at the nonsouthern ends of these three flows should develop labor surpluses, capital shortages, and money outflows. These developments should, in turn, lead to lower nonsouthern wages, lack of investment income, and leakage of existing income. In short, nonsouthern real income should go down while southern income is going up. This analysis oversimplifies, but it does show how traditional theory calls for, and might explain, income convergence.

The trouble is, recent American economic history conclusively refutes the idea that traditional theory can explain income convergence. Despite some elements of truth, the theory proves to be predominantly false in all three of its aspects: labor flows, capital flows, and commodity flows. Labor, capital, and commodities have not flowed in ways that would cause income to converge.

Labor Flows

The theory says labor flows will equalize regional wage rates. In other words, labor should flow out of the low-wage South to other regions, creating labor shortages in the South and driving up wages. But that notion comes under immediate suspicion. Labor flows were often in the "wrong" direction, and southern wages didn't equalize. In 1965, or 100 years after the Civil War, the Deep South's average hourly wage in manufacturing was still only 76 percent of the nonsouthern average. We use the Civil War as a point of reference because that is when the ceiling that slavery placed over southern wages was lifted; hourly labor no longer competed against unpaid labor, so wages became free to climb.

The failure of the labor flow mechanism to equalize wages points to four problems with the labor flow aspect of traditional theory: (1) industry mix and other mix effects, (2) job availability effects, (3) amenity effects, and (4) sharecropper reserve labor effects. Take economic mix. The concept of a regional wage rate oversimplifies. There is no single wage rate; there is a wage structure, analogous to the interest rate structure. Wages differ by industry, occupation, rural-metropolitan location, and so on. Even if labor flows were to suddenly equalize the regional wage rates for specific industries and occupations, substantial differences in average wages would remain. Average wage rates would remain low in states whose employment mix was heavily weighted with low-wage manufacturing industries, such as textiles and apparel; agriculture; nonprofessional occupations; and nonmetropolitan job locations. The South is a case in point. The South's employment mix *is* heavily weighted with textile and apparel manufacturing, agriculture, nonprofessional occupations, and nonmetropolitan job locations. The South also has the nation's highest black population percentage, with blacks generally earning lower wages than whites.

An even worse problem with traditional labor flow theory is that regional wage

differences have little effect on migration. What counts is jobs. Job availability, not wage rates, is the commanding influence in migration. Unemployed workers leave high-wage states (the Manufacturing Belt) and average-wage states (the Plains) where jobs are scarce, because a job is far more important than a good wage. Many of these workers go to the nation's low-wage region, the South, for the same reason: the South has rapid employment growth, and a job is more important than a high wage. Consider these facts:[2]

- In the Manufacturing Belt, industrial decline—lack of jobs—produced net out-migration in the 1970s and 1980s (through 1986), despite the region's having (by 1970) the nation's highest wage levels. (What good is a high wage if you can't get the job that pays it?) The Manufacturing Belt did have net in-migration in the 1950s and 1960s, during the sharecropper exodus from the South, but even then the region had net out-migration among whites.
- In the Plains, agricultural decline—lack of jobs—produced net out-migration in every postwar decade, despite wages that were essentially at (97 to 98 percent of) the national average.
- In the South, rapid industrial growth—abundant jobs—produced net in-migration in the 1960s, 1970s, and 1980s (through 1986), despite the region's having the nation's lowest wage levels. In the 1950s, at the height of the sharecropper exodus, the South did have net out-migration, but only in two of its three census divisions; the South Atlantic division had net in-migration of 3.1 percent.
- In the West, rapid industrial growth—abundant jobs—produced net in-migration in every decade. High wages in the West might have been a supplemental factor in this region, but probably not: during 1975-80, 2.7 people migrated from the Manufacturing Belt to the West for every person who went in the reverse direction, despite slightly higher wages in the Manufacturing Belt.[3]

A third problem with the labor-flows-to-high-wage-regions theory is that it ignores amenity effects. Noneconomic forces like amenities tend to be ignored in neoclassical economic analysis, which likes to reduce everything to supply, demand, cost, and price. Yet migration studies have shown that some migration is influenced by such amenities as mild winters, seacoasts, lakes, and mountains.[4] Why does labor flow out of the Manufacturing Belt despite the region's having the nation's highest wage levels? The main reason, we have seen, is lack of jobs. But a secondary reason seems to be a relative lack of amenities. True, the Manufacturing Belt does have amenities, but it is not as well endowed as the South and West. Climate is particularly important in this respect: the Manufacturing Belt has harsh winters.

The fourth problem with labor flow theory is that it overlooks the effects the sharecropper-tenant farmer system had on the South's labor supply. Sharecropping was an extremely labor-intensive form of agriculture. It gave the South a huge supply of reserve labor—labor that could shift to the industrial sector (and the urban sector) whenever incipient labor shortages threatened to drive wages up. In the South, urban wages were *much* higher than the crop-share wages

of sharecropping, so there was never a lack of new workers to fill job vacancies in southern industry. The South's reserve labor—the sharecroppers—prevented the emergence of the labor shortages that labor flow theory predicts.

Between 1950 and 1970, the reserve labor pool dried up as the sharecroppers left their farms. Only then, as Browne observes, did southern wage gains contribute appreciably to the South's per-capita income gains.[5] Moreover, the labor shortages that caused the wage gains could not be attributed to low southern wages, at least not in an important degree. The plain truth is that new technology, not low wages, is the main force that pushed the sharecroppers off the farms. They had to leave, because there was no longer any work for them. And the millions who left the South did so not so much because of the lure of northern wages (though that was an appreciable influence) but because discrimination made it hard for them to find decent jobs in the South.

Capital Flows

The second of the three neoclassical flows affecting income is capital flows. Peeling away the abstraction ("capital"), we come to new manufacturing plants as the essential ingredient in this flow. The idea is that new plants flow to (locate in) the places that offer the lowest costs. The low costs could apply to either labor or raw materials. For manufacturing in general and for raw materials in general, the South has no appreciable raw material cost advantages over other regions. The region does have good timber resources; and petroleum output in Louisiana, Oklahoma, and Texas gives the South a cost advantage for petroleum refining and petrochemicals. But in the main, any capital flow effect on southern income must result from the South's labor cost advantage. The South has the nation's lowest hourly wages. It also has the weakest unions. Unions can add to labor costs through work rules, strikes, slowdowns, militant attitudes, and resistance to innovation.

The question is: have low labor costs in the South attracted manufacturing (capital), thereby stimulating southern income and causing convergence? In the main, the answer is no. Low labor costs have definitely attracted *some* industry to the South. But only 3 or 4 of the 20 two-digit manufacturing industries are strongly concerned about labor factors; the rest are predominantly market-oriented.[6] The labor-oriented industries are textiles, apparel, leather goods, and to a limited extent furniture. (Furniture is primarily market-oriented, but upholstered furniture is attracted to textiles, which is attracted to low labor costs.)

Consequently, low labor costs cannot give and have not given the South the nation's highest manufacturing growth rates. Studies of three periods—1947–63, 1963–77, and 1976–84—show that the West, not the South, had the highest manufacturing growth rate in each period.[7] For 1976-84, the West's 24.7 percent manufacturing growth rate far exceeded those of 11.1 percent for the Southeast, 8.7 percent for the South Central states, 4.7 percent for the Plains, and -7.9 percent for the Manufacturing Belt.[8]

The 1947-63 and 1963-77 studies provide comparisons of the seven leading industrial location determinants. These comparisons show how much of the

period's variation in state percentage increase in manufacturing employment is explained by each determinant. In both periods, six factors explain 96 percent of the variance. (One factor differs between periods, which is why there are seven factors altogether.) Table 3.1 shows how the seven factors rank in importance. In both periods, markets is the dominant factor, explaining nine or ten times as much of the variance as labor does. Markets explains 60 percent of the 1947–63 variance and 55 percent of the 1963–77 variance. Labor explains only 6 percent of the 1947–63 variance and, likewise, only 6 percent of the 1963–77 variance. Incidentally, the labor finding shows that the sclerosis theory, discussed in the next section, grossly overstates the effect of unions on manufacturing growth.

Many empirical studies affirm the dominance of the market influence.[9] To understand why the labor influence is so weak, we must understand why the market influence is so strong. Manufacturing has always been concentrated in the Northeast-Great Lakes Manufacturing Belt. But since at least 1900, and particularly since World War II, it has been shifting (not migrating) to the South and West.[10] Regional branch plants producing for regional (short-haul) distribution are taking over from national plants producing for national (long-haul) distribution; the factory is getting closer to the market. The underlying reason for this change is a desire to reduce factory-to-market shipping costs by putting new production closer to the undersupplied markets of the South and West. Both the South and the West, not to mention the Plains, offer the desired transportation cost savings. But the West offers the biggest savings, for two reasons. First, the western states are farther from the Manufacturing Belt than are the states of the Plains and the South; it costs more to ship goods from the Manufacturing Belt to the West than to the other undersupplied regions. Second, the West has the lowest manufacturing employment per capita, so proportionately more of the goods sold in the West must be shipped from the Manufacturing Belt.[11] Because the West offers the biggest savings, it attracts the most new branch plants. In a related vein, *new* firms locating in the West are protected from Manufacturing Belt competition by a relatively high transport cost "tariff."

We see, then, that transport cost savings, which favor the West, far outweigh labor cost savings, which favor the South. For this reason, traditional theory's capital flow mechanism is largely impotent. It fails to channel the biggest capital flows to the low-income region, the South; indeed, it channels the biggest capital flows to 1950's *high*-income region, the West.

Of course, the traditional theory can be interpreted more broadly to include distribution (transportation) costs as well as production costs among the costs that determine rates of return, hence capital flows. But in this case, the theory calls for manufacturing growth to be fastest in the West rather than the South: the West offers the biggest transport cost savings. To be sure, the South offers a transport cost advantage (equivalent to a high rate of return on investment) vis-à-vis the Manufacturing Belt, but the West has an even bigger advantage. The theory thus implies that manufacturing growth should be high in the South but highest in the West. And as we have seen, what should be is actually the case.

The trouble is, this outcome further implies that the per-capita income gap between the West and the South should *widen*, even while the gap between the South and the Manufacturing Belt narrows. But the West-South income gap has

Table 3.1
Percentage of the Variation in Manufacturing Employment Growth Rates Explained by Significant Factors, 1947–1963 and 1963–1977

Factor	1947–63	1963–77
Markets	60	55
Climate	23	15
Rural Attraction	0	11
Urban Attraction	2	0
Labor	6	6
Thresholds	5	5
Retiree-Amenities	not tested	4
Percentage Explained	96	96

Source: Leonard F. Wheat, "The Determinants of 1963–77 Regional Manufacturing Growth: Why the South and West Grow," *Journal of Regional Science* 26, no. 4 (1986), table 3, 651.

not widened; it has narrowed. Moreover, the West's 1950 per-capita income advantage over the Manufacturing Belt has reversed: the Manufacturing Belt now has a slight advantage over the West, despite the West's extremely high manufacturing growth rate. Traditional theory's capital flows obviously can't explain what has been happening.

Commodity Flows

The third of traditional theory's three flow mechanisms—labor, capital, and commodities—involves commodity flows. In theory, the region with the lowest costs should have the lowest prices for goods. The low-cost region's price advantage will attract buyers and cause commodities to flow from the low-cost region to high-cost regions. Applied to the South, the low-labor-cost region, the theory suggests that exports from the South to other regions should stimulate income growth in the South, producing convergence.

But what the theory predicts has not happened to any appreciable extent; the South has not become—and is not becoming—a major net exporter of goods to other regions. Granted, market-oriented manufacturing growth has made the South more industrially self-sufficient. That is, import substitution has occurred, and this has certainly helped the South's income position. And in particular industries—mainly textiles, apparel, furniture, tobacco, petroleum refining, and petrochemicals—the South does export more than it imports. Yet even most of these exceptional industries have considerable nonsouthern production; apparel and furniture manufacturing can be found all over the nation.

Why has the commodity flow mechanism provided only modest help in closing the southern-nonsouthern income gap? The answer is that low production costs won't generate heavy exports if they are offset by high distribution costs. To repeat, most industry is market-oriented. Most new plants are branch plants producing for regional markets. Lower *delivered* costs usually result when goods are produced regionally, even if this means slightly higher *production* costs. Low distribution costs on "home-grown" goods in the nonsouthern regions offset higher production costs. This cost trade-off explains why the West is building its own factories instead of importing goods from the South. A related consideration is that most industries are fairly capital-intensive; the labor cost savings offered by the South are relatively small in the overall cost picture. For these reasons, traditional theory's commodity flows will not equalize regional incomes.

THE INSTITUTIONAL SCLEROSIS THEORY

Mancur Olson, in *The Rise and Decline of Nations* (1982), develops a theory that tries to explain why some nations and subnational regions grow faster than others. Olson seeks to explain growth in both total income and per-capita income. His theory holds that special interest organizations and institutions impede growth. As nations and regions grow older, their interest groups grow stronger and more numerous. These groups do various things that interfere with growth. As a result, older societies develop a condition Olson dubs "institutional sclerosis."[12] The theory says sclerotic economies are largely responsible for the relatively slow growth found in (*a*) nations with well-developed groups and institutions and (*b*) the older states and regions of the United States.

This is no narrow theory about trade associations and unions. The interest groups include all those that use lobbying and other pressure group tactics. In other words, Olson is referring to the whole spectrum of interest groups—the churches, the Sierra Club, the United Auto Workers, the National Rifle Association, Mothers Against Drunk Drivers, the National Glass Association, Common Cause, Planned Parenthood, the National Organization for Women, and all the rest. All cause sclerosis.

In our judgment, the sclerosis theory is wrong. It is badly—and vaguely—reasoned, largely ignores many other factors that genuinely explain fast and slow growth, and relies on gravely defective econometric tests.

Faulty Causal Mechanisms

Searching Olson's text for causal mechanisms that explain the supposed effects of interest groups, one is struck by the abstractness of the analysis. The theory lacks coherent, detailed, carefully worked out causal mechanisms. The road from cause to effect has unbridged canyons; we get hints where clear explanations and clearly stated premises are needed; the explanations we do get are superficial and raise more questions than they answer. At times the theory seems to imply the absurd: that interest groups make or control decisions on where to put new branch

plants. These defects surface when we inspect the five overlapping causal mechanisms found scattered about the text:

1. A particular set of groups, unions, impedes growth in many ways.
2. Interest groups make decisions too slowly.
3. Interest groups delay the introduction of new technology and innovation.
4. Interest groups lobby for protectionist legislation and bail-outs.
5. Interest groups lobby for regulation; regulation hampers growth.

Now let us see why these causal mechanisms fail to substantiate the theory.

Unions

The unions-impede-growth mechanism is not really separate; it cuts across the other four. But Olson draws so heavily on unions for examples that the essence of a separate causal mechanism emerges. The theory keeps coming back to unions. They do this, that, and the other thing that stymies growth. And in Olson's statistical tests, a unionization variable is used to measure the strength of all *other* interest groups.

It is certainly true that unions discourage growth. Many studies—most are older than Olson's—have found that industry tends to avoid unions.[13] The reasons are obvious and well understood. They include strikes, slowdowns, grievances, work rules, other perceived interference with "management prerogatives," resistance to labor-saving innovations, featherbedding, and high wages.

But to generalize from unions to all interest groups and institutions—to imply that other groups have similar negative effects—is illogical. The relationship of unions to manufacturing is unique. No other interest group even approaches that relationship. Union members *work* for manufacturing firms. Through strikes they can painfully damage—even destroy—firms. Through work rules, slowdowns, and high wages, they can undermine productivity and raise costs. What other group has comparable power vis-à-vis manufacturing? Certainly not the American Legion, the National Peanut Council, Public Citizen, the National Abortion Rights Action League, or the Air Transport Association.

The truth is, unions are a case in themselves—an independent influence on growth. Yet Olson in effect argues that, because unions deter growth, therefore the National Organization for Women and all other interest groups deter growth. Such reasoning is fallacious. One cannot generalize from unions to all interest groups; union effects are not evidence that nonunion groups have similar affects. Olson must show that other organizations have their own effects. Only where the other groups are concerned does Olson have a new theory.

This last point deserves elaboration. Olson cannot claim as his own the theory that unions discourage new manufacturing from locating in places where unions are strong. Nine years before the sclerosis theory was published, one of us wrote:

> A very *popular* [emphasis added] hypothesis is that labor attractions—low wages, weak unions, and surplus labor—stimulate manufacturing growth, most notably in the South. . . . Industry's aversion to unions requires little elaboration. The threat of strikes and other work stoppages, disputes about "management prerogatives," and the efficacy of unions in raising wages makes a union anything from a nuisance to an abomination in the eyes of most companies (not that firms necessarily seek to avoid unions).[14]

The only thing original or novel in the sclerosis theory is the idea that *nonunion* interest groups impede growth. That idea is the essence of the theory. No amount of evidence that unions impede growth can validate the theory.

Slow Decisions

The second causal mechanism by which groups allegedly impede growth is slow decision making by interest groups. (This mechanism overlaps the third.) Just how slow decisions cause slow growth is never spelled out. What *is* clear is the amazing *implication* that interest groups effectively control the decisions of manufacturing firms to innovate and to build new plants. This isn't entirely implication: slow group decisions explicitly affect "the rate of adoption of new technologies."[15] (Is Olson still thinking about unions, or does he really mean to imply that the Save the Redwoods League and the Knights of Columbus hinder the adoption of new manufacturing technologies?)

The slow-decisions causal mechanism has four defects. First, despite all the theorizing about "crowded agendas" and "consensual bargaining" among group members, the fact is that interest groups generally decide things quickly.[16] Unions hold a meeting, take a vote, and strike. Other groups, responding to perceived threats to their interests, quickly decide to place newspaper ads, organize rallies, and mobilize their lobbyists. They can act quickly because, like corporations and governments, they have bureaucracies and executives to make decisions. The executives have broad authority and operate under well-established general policies. These general policies mean that, in effect, many decisions are made in advance and merely need to be implemented in specific situations as those situations arise.

Second, even if the decisions were slow, that would be irrelevant: interest groups are not making the decisions that affect growth. The critical decisions are the decisions of *firms*—not cartels—to (*a*) build new branch plants or otherwise expand output and (*b*) adopt new technology. We find in Olson's theory considerable loose talk about "cartels" and "collusion"—but hardly any real evidence or concrete examples outside of unions. From the theory, one would think that most trade associations are cartels. One would think that General Motors had to obtain (through "consensual bargaining") the approval of the Motor Vehicle Manufacturers Association before building its innovative Saturn plant in Tennessee. Such ideas embody serious misunderstanding of how American industry operates. Firms—not groups—make the decisions to build new plants. And having made their decisions, the firms need not and do not sit around and

wait for their industry association or the Daughters of the American Revolution to slowly decide to ratify the new-plant decisions.

Third, even if interest groups made the decisions on new plants, "their slowness in making decisions" would not cause a slow "rate of growth."[17] The slow decisions would cause only a lower *level* of cumulative growth. Lengthening the time it takes to make decisions is like lengthening a pipe through which water flows: the water spends more time in the pipeline, but it flows out the far end just as fast. Likewise, if new plants spent more time in the planning pipeline, this would not affect the number of new plants built each year.

Fourth, even if interest groups did slow the rate of plant formation, that would not explain American's regional growth rate differences. The speed with which new plants come into existence does not determine *where* those plants go. If General Motors decides that Tennessee is the best place for its Saturn plant, that is where the plant will go—regardless of whether some cartel delays the plant for two years.

New Technology and Innovations

Interest groups block new technology and innovation, according to the theory, because (*a*) unions oppose changes, (*b*) slow decisions by interest groups delay innovations, and (*c*) "colluding firms," such as trucking firms and airlines, sometimes get "monopoly rights" that innovations would cause them to lose.[18]

Let's begin with the three reasons. The first two we have just discussed. Reason three—collusion and monopoly—alludes to cases where a single common carrier was certificated by the Interstate Commerce Commission or the former Civil Aeronautics Board to serve a particular city. Here Olson confuses regulatory decisions (sometimes—in past decades—designed to minimize airline subsidies) with cartel decisions. You might call this the devil theory of single-carrier service, the devil being "colluding firms." Olson does not say what innovations were blocked, but he seems to have new equipment in mind. The idea seems to be that competition would lead to the use of more modern planes and better trucks at points now served by one carrier. Olson says "modern American economic history" offers "many examples," but he doesn't give even one.[19] This sort of evidence—a mere assertion that there *is* evidence—is not acceptable.

It is at least as likely that interest groups spur innovation as it is that they delay it. Unions? Under John L. Lewis, the United Mine Workers won high wages. The high cost of labor in the coal mining industry forced the introduction of labor-saving mining machines. High wages also pushed up the price of coal. That development accelerated the introduction of a railroad innovation: diesel locomotives. Collusion by truckers? The American Trucking Association has successfully lobbied for laws that permit the use of labor-saving double-bottom (dual trailer) tractor-trailers. Ralph Nader's consumer interest group, Public Citizen, has lobbied for regulations requiring automobile innovations—seat belts and energy-absorbing steering columns, for example. Timber groups have lobbied for clear-cutting, a sorrowful but nevertheless cost-saving technology. Action on Smoking and Health has pressured airlines into banning smoking on most flights.

That innovation will reportedly save the airlines tens of thousands of dollars formerly spent cleaning tar-clogged ventilation equipment.

Olson argues that small interest groups, such as trade associations in oligopolistic industries, can exercise their power more easily. So be it. But that very smallness also tends to encourage innovation: "cartels" beget innovation. John Kenneth Galbraith observes that price competition in oligopolistic industries is generally painful and unprofitable. This condition diverts competition into another channel—innovation. Galbraith acknowledges that patent suppression has occurred in some exceptional cases. But, he observes:

> To maintain a convention against innovation requires a remarkably comprehensive form of collusion. This is difficult as well as legally dangerous. While it would be going too far to say that oligopoly insures progress, technical development is all but certain to be one of the instruments of commercial rivalry when the number of firms is small.[20]

He adds that "technical development [contrasted with price cuts] is a safe rather than a reciprocally destructive method by which any one firm can advance itself against its few powerful rivals."[21]

The delayed-innovation argument has one other flaw. The trade associations that supposedly impede innovation are national organizations. If they block an innovation, they will block it nationwide. The nation as a whole will then suffer. But how could that explain the faster growth in the South and West vis-à-vis the Manufacturing Belt?

Protectionist Measures

The fourth causal mechanism by which groups impede growth, according to Olson, is protectionist legislation. The theory says that interest groups foster protectionism. "One obvious way in which they do so is by lobbying for bail-outs of failing firms."[22] Protectionism delays the reallocation of resources to better uses.

This argument is well calculated to appeal to economists. Unfortunately, the argument does not fit the theory as applied to American states or regions rather than to nations. Protectionist measures are national measures, not state measures. We do not have state tariffs and state import restrictions. And since the national measures apply equally to all states, they do not explain why the South and West grow faster than the Manufacturing Belt.

Granted, one might argue that industry mix differences give particular tariffs more effect in certain states. But the injury (slower growth) is not necessarily in the more protected states. It is in whatever states would have attracted the alternative industries if resources had been permitted to flow from protected industries to unprotected (or less protected) ones.

Furthermore, the protectionism argument begs the question: it assumes that interest groups rather than firms lobby for bail-outs. If groups—trade

associations—are lobbying for bail-outs, healthy firms are collectively working to save their sick competitors. Did the rest of the auto industry try to save Chrysler and American Motors when those firms were failing? Or is Olson confusing lobbying by *firms* with lobbying by interest groups?

Lobbying for Laws and Regulations

The fifth way interest groups supposedly hurt growth is by lobbying for laws and regulations. Groups "increase regulation" through "lobbying and the complexity of legislation it brings about." The regulations diminish "the incentive to produce" and increase "the incentive to seek a larger share of what is produced."[23]

An obvious flaw in this argument is that it contradicts itself. How can a firm gain "a larger share" of its industry's output without increasing its own production? Doesn't the incentive to seek a larger share imply an "incentive to produce" *more*? Granted, a firm could also get a larger share by holding production constant in the face of falling industry output, but that obviously is not what Olson has in mind. The whole point of getting a larger share is to increase profits by increasing sales and, hence, production.

Olson seems to realize that this argument is not only weak but dangerous to his theory: he never really develops the argument or pushes it very hard. The argument is dangerous because it relegates interest groups to a minor role. Regulation becomes the real culprit. But regulatory laws and supporting administrative regulations arise from social and economic problems. These problems are rooted in the complexity of advanced societies and economies. The notion that regulation results from lobbying by interest groups is simplistic. We have regulations not because of interest groups but because the problems regulations attack are viewed as serious enough to warrant regulation.

The argument that regulation causes slow growth is also dangerous because it can be turned against the sclerosis theory. True, organizations like Common Cause, the Center for Auto Safety, and the Save Lake Superior Association lobby for regulation. But other groups lobby *against* regulation. The auto industry lobbied against mandatory seat belts and energy-absorbing bumpers. The timber industry continues to lobby against bans on clear-cutting. And many industries have lobbied against pollution controls. David Truman and John Galbraith have pointed out that interest groups lobby on both (if not three or four) sides of every issue.[24] When it comes to regulation, surely the groups that oppose regulation outweigh the groups that support it. When the pro-regulation forces win, it is generally because the weaker lobbying groups have gained strong public support—support that is often based on some well-publicized incident or situation. Thus, *if regulation hampers growth, growth should be fastest in states where strong interest groups can block regulation.*

Apart from these flaws in his argument, Olson has the further problem of showing that regulation actually does inhibit growth. Stafford found that environmental regulations have little effect on growth.[25]

Other Problems With Sclerosis

Beyond faulty causal mechanisms, the theory almost totally ignores—and shows serious misconceptions about—the real causes of growth differences. It also uses gravely flawed econometric tests. For brevity's sake, we will merely summarize the main considerations behind these two points. More detailed analysis can be found in a paper by Wheat.[26]

The Real Causes of Growth Differences

If sclerosis does not explain the growth rate differences among states and regions of the United States, what does? To answer this question, we must distinguish between two types of growth that Olson discusses: (1) size growth and (2) per-capita income growth. Size growth is growth in population, employment, income, and the like; it is rooted in physical growth in population. To Olson, size growth and per-capita income growth are simply alternative measures of the same basic phenomenon. He believes variation in both types of growth has the same principal cause, sclerosis. And because both have the same cause, both types of growth implicitly have the same geographic pattern.

Unfortunately for the theory, size growth and per-capita income growth have considerably different geographic patterns. True, both are high in the South. But consider these contrasts:

1. Size growth is highest in the West, but per-capita income growth is *lowest* in the West.
2. Size growth is second-lowest (relatively weak) in the Plains, but per-capita income growth is second-*highest* (relatively strong) in the Plains.
3. Size growth is lowest in the Manufacturing Belt, but per-capita income growth is lowest at the opposite end of the country—in the West.

The observed geographic patterns conflict with the idea that weak interest groups are the main force behind growth. According to Olson, interest groups are strongest in the Manufacturing Belt. They are weakest in the South and in the newest states, those of the West. Olson leans toward the view that the very weakest groups are in the South: two of three adjusted time variables he uses to measure group strength treat the South as having the weakest groups, and so do some lawyers-per-capita and unionization variables he also uses.

It is certainly true that per-capita income growth is highest in the South. But if interest groups are strongest in the Manufacturing Belt, why is per-capita income growth lowest in the West, where interest groups supposedly are weak (and possibly weakest)? And if interest groups are weakest in the South, why is size growth highest in the West? Also, if size growth and per-capita income growth have the same principal cause, sclerosis, why is size growth highest in the West whereas per-capita income growth is lowest?

The answers to these questions are that (*a*) size growth and per-capita income growth do not have the same causes and (*b*) neither type of growth has anything to do with institutional sclerosis. Moreover, and contrary to what Olson seems to assume, per-capita income is not an important determinant of total income (size). Total income (Y) is determined almost entirely by total population (P): the 1980 correlation between Y and P for the 48 contiguous states is +.9935. Although Y = P x Y/P, where Y/P is per-capita income, P displays extreme variation, whereas Y/P has a comparatively narrow range of variation. The 1980 population of California (highest) was 50.4 times as high as that of Wyoming (lowest of 48), but the 1980 per-capita income of Connecticut (highest) was only 1.7 times that of Mississippi (lowest). Because P varies so much more than Y/P, P overwhelms Y/P when Y/P is multiplied by P to get Y.

If size growth (and decline) and per-capita income growth have different causes, what are those causes? The final section of this chapter discusses the causes of per-capita income change; they have been previewed in chapter 1. At this point we will simply note that the sclerosis theory shows no awareness of any of the influences that underlie per-capita income change, except for manufacturing growth (a minor influence). As for the causes of size growth, the analysis can be simplified to an analysis of manufacturing growth. Changes in other basic industries—agriculture, the extractive industries, government—also influence size growth, as does non-job-related migration, especially among retirees. But manufacturing change is the primary determinant of size growth for most states. And, equally important, manufacturing is the principal medium through which sclerosis supposedly operates.

Manufacturing's growth pattern has been one of stagnation and decline in the Northeast-Great Lakes Manufacturing Belt, moderate growth in the Plains, fast growth in the South, and even faster growth in the West.[27] Growth in population (P) and income (Y, or P x Y/P) has followed the same relative pattern, except that agricultural decline in the Plains has caused slower population growth (and occasional decline) in that region.

Earlier in this chapter we described the locational importance of the market influence, which is causing industry to decentralize from the Manufacturing Belt to the South and West. Empirical research since World War II provides a strong consensus that markets are the dominant influence behind regional manufacturing growth.[28] The market influence is the tendency for supply (manufacturing employment) to become geographically prorated according to demand (population and income). Industrial location theory supports the empirical research. Location theory holds that the leading plant location influence is markets, sometimes conceptualized as transportation, or factory-to-market shipping distance, time, and cost. Surprisingly, the sclerosis theory does not recognize the well-known, widely discussed market influence and its role in the relatively fast manufacturing growth of the South and West.

The power of the market magnet registers in table 3.1, discussed earlier in connection with flaws in neoclassical theory. Table 3.1 has findings from two studies of manufacturing growth rates—a 1947–63 study and a 1963–77 study.[29] The market influence, measured by demand/supply ratios, is stronger than all other influences combined. It explains 60 percent of the manufacturing growth

rate variation for 1947–63 and 55 percent for 1963–77. Two other influences—climate and amenities—are related to markets: mild winters and amenities attract migrants, and the resulting population growth leads to the growth of certain local-market-oriented industries. If this climate-amenity-based growth is credited to markets, the market influence explains 75 percent of the 1963–77 variation in growth.

The manufacturing growth influences additionally include a 1947–63 urban state attraction involving farm-to-metropolitan migration, a 1963–77 rural state attraction involving rural labor and noncompetitive markets, the negative influence of unions, and a threshold influence related to agglomeration economies. The full set of influences explains 96 percent of the growth rate variance for both periods. The 4 percent that is unexplained is probably attributable to a combination of (*a*) the resource influence, which cannot be tested adequately in aggregate industry analysis, (*b*) productivity increases and foreign competition, which caused employment declines in textiles, apparel, automobiles, steel, and other industries, and (*c*) imperfect measurement of curvilinear regression. There is no reason to invoke sclerosis or to believe that a sclerosis influence exists.

Econometric Flaws

The flaws in Olson's econometric tests are extensive. One problem is that none of the four variables he uses to measure the strength of interest groups really measures that. His main variable, *years since statehood*, is actually a proxy for markets, which are strongest in the newest states—the western states. Indeed, 20 years before Olson used the statehood variable, Fuchs used it in a manufacturing growth study to measure the market influence![30] Fuchs was testing the hypothesis that the less developed states are "catching up" in their relative levels of manufacturing. Another two of Olson's variables—*lawyers per capita* and *1880 urban population percentage*—have high values in metropolitan areas and tend to identify the highly metropolitan Manufacturing Belt states. These states grow slowly not because of strong interest groups but because of weak markets, high wages, heavy unionization, and generally harsh winters. Olson's other variable is *extent of unionization*. But whereas Olson uses this variable as a proxy for nonunion interest groups, the obvious truth is that the variable speaks for itself: it describes the negative influence of unions on manufacturing growth.

Another major flaw is that Olson's main variable, years since statehood, uses adjusted statehood dates for the southern states—old states that, contrary to the theory, have weak unions and fast growth. In three alternative statehood variables, the old states become new states. The Confederate states get revised statehood dates of 1865, 1915, and 1965. In each case the variable is 1965 minus statehood year. The 1965 version thus gives the southern states zero years since statehood. This is like creating a dummy variable valued at one for nonsouthern states and zero for southern ones. Such a variable can register the fast growth of southern states (a negative correlation), but it cannot identify sclerosis or anything else as the cause of that growth.

Olson's third grave error is failure to hold other variables constant. Control—holding constant the other things that could cause variation in the dependent

variable (growth)—is perhaps the most fundamental requirement of hypothesis testing. Yet except in a few equations that include a catch-up variable, Olson's regression equations have no independent variables other than the misnamed sclerosis variables. The catch-up variable, state deviation from national per-capita income, "decisively outperforms the [sclerosis] variables" in Olson's per-capita income growth equations.[31] And beyond that problem, the catch-up variable does not even belong in the equations, because it registers an effect (catching up by the low-income states) rather than a cause of growth. Ultimately, then, Olson has still failed to control the real forces behind income growth—changes in agriculture, education, employment/population ratios, racial mix, western transportation costs, and so on.

AN EMPIRICALLY BASED THEORY OF CONVERGENCE

If neither Heckscher-Ohlin type neoclassical theory nor Olson's sclerosis theory can explain per-capita income convergence, what *is* the explanation? There is an abstract reason for convergence, and behind it are many concrete reasons. We can best understand these reasons—abstract and concrete—by looking separately at (1) the South, (2) the nonsouthern regions, and (3) the overall United States.

The South

Abstractly speaking, incomes converged because the regions became more alike—in racial mix, farm-urban mix, industry mix, occupational mix, the capital intensiveness of agriculture, educational attainment, per-capita employment, cost of living, and wage levels. In all these respects, the South has always had an income disadvantage vis-à-vis all other regions. (The South's low cost of living is an advantage in one sense, but from a per-capita income standpoint a low cost of living encourages low wages and thereby lowers income.) The combined effect of so many disadvantages has been southern per-capita income levels well below those of the rest of the country.

But since 1950, the South has declined in black population percentage, urbanized, industrialized, become more capital-intensive in its agriculture, greatly improved its educational levels, gained in per-capita employment, and developed labor shortages that put upward pressure on wages. Where rural-urban mix, industry mix, and occupational mix are concerned, partly similar developments have occurred in the heavily agricultural Plains (West North Central) region and in agricultural states elsewhere: the technological revolution in agriculture and the consequent off-the-farm exodus have helped raise per-capita income in comparatively rural, agricultural states everywhere, not just in the South. But because conditions in the South were more extreme, and because the South had more disadvantages, the South's per-capita income gains were greatest.

The South's central role in the income developments since 1950 points to a unifying theme. More than anything else, the income convergence of 1950-87 was a matter of the South's overcoming the legacy of slavery. The Civil War left the

South with a shattered economic system. Slavery and the plantation system were dead. A new economic system was needed to replace the old. The new system that evolved was the sharecropper-tenant farmer system, ably described by Persky, Wright, and Lemann (summarized in chapter 2). The planters could not afford to pay cash wages, at least not in the early years, so labor was paid in kind. The sharecroppers and tenants received a share of the crop; they had little cash income, sometimes none. At first the sharecroppers and tenants were freed slaves, but over time the system became biracial. The new system's result was a huge poverty class—the sharecroppers and tenant farmers—not found in any other region. And outside of agriculture, the South's urban blacks were also impoverished. These poverty populations, part of slavery's legacy, were the leading cause of the South's low per-capita income.

But there was more to slavery's legacy than sharecroppers, tenant farmers, and impoverished urban blacks. Slavery also left the South with the nation's most agricultural economy, its most rural society, its worst educational systems, its lowest labor force participation rates, and its lowest wage rates. The plantation system revolved around agriculture and discouraged industry. Maddox observes that "the slave-based system of plantation farming . . . brought into being an agrarian pseudoaristocratic form of social organization . . . and provided an environment basically hostile to technological innovation and social change."[32] Wright finds that "the closed character of the southern population was directly related to the limited development of manufacturing in the slave economy."[33] The South thus emerged from the Civil War with an unfavorable rural-urban and industrial mix, heavy on agriculture (low income) and light on manufacturing (high income).

Slavery and its attendant aristocratic society also gave the South a poor educational system. It is common knowledge that racial prejudice and segregated schools imposed severe educational handicaps on blacks. Less understood is the fact that poor whites also suffered. The planters, who held political power, believed that education was a ticket off the plantation. They believed that workers who could read and write and do arithmetic would not be satisfied with field work; educated workers would seek jobs in a town or city. The planters also viewed education as something that would encourage workers to leave the South: the region would not reap the rewards of its education investment. "For reasons such as these, the South's investment in education was far below [that of] the rest of the nation."[34]

Low labor force participation rates were also part of the legacy of slavery. The percentage of the population that enters the labor force depends on urbanization, industrialization, education, and the percentage of the population that is of working age. Urbanization and industrialization create jobs for women. Education opens employment doors. Working-age people tend to work. Slavery gave the South an agricultural, nonindustrial, poorly educated socioeconomy. Slavery thus contributed to low participation. High fertility among rural families, meanwhile, leads to high population percentages of children and consequently low percentages of working-age persons. Low participation in the South meant that relatively fewer persons were earning incomes.

As for wages, slaves don't get paid. It is true that slaves sometimes were

rented out; they sometimes commanded the equivalent of a wage. But the fact remains that white labor competed against unpaid and low-rent black labor. The result was extremely low wage scales. Thus, because of slavery, the South emerged from the Civil War with the lowest wage scales in the nation. And it apparently takes well over 100 years for economic forces to bring such low wages up to the national average.

The multifaceted legacy of slavery provided the basis for the post-1950 developments that caused southern income to close in on the national average. First and predominant among these developments was the collapse of the sharecropper-tenant farmer system ("sharecropping" for short). The causes of sharecropping's collapse included mechanized harvesting (the cotton picker especially), mechanized cultivation (the spread of tractors), chemical control of weeds (herbicides), and an agricultural shift from cotton to less labor-intensive farm products. The pull of better-paying jobs in the South's growing industrial sector—and in the North—was also a factor.

Sharecropping's collapse had pervasive effects on income. One effect was an interregional transfer of poverty. As already noted, 5 million blacks migrated to other regions; so did uncounted numbers of poor white former tenants. The result: the South had proportionately fewer poor—extraordinarily poor—people than before, whereas the North had proportionately more. Moreover, the people who remained on southern farms enjoyed higher incomes. The new agriculture was more capital-intensive and displayed bigger (consolidated) farms. Hence the farm workers tended to be equipment operators (better paid) rather than manual laborers, and the farm operators had income from more acreage.

The off-the-farm migration, in both its interregional and intraregional aspects, also changed the farm–urban balance of the southern economy. The South became less agricultural (low wage) and more urban (high wage). This development can also be viewed as a change in industry mix and occupational mix. Cities, especially metropolitan ones, have more favorable industrial and occupational mixes. They have more employment in manufacturing, transportation, finance, and many other industrial sectors that pay better than farming. Cities also have more professional, managerial, technical, and skilled trades employment. Somewhat related to these mix changes was a change in educational attainment. Urban residents tend to be better educated, hence better equipped to earn good incomes; this tendency was especially strong in the South. Although the adult migrants received no educational benefits from migration, their children found better school systems. Partly for this reason, educational attainment climbed in the South.

The collapse of sharecropping was the first post-1950 development in the South. The second was the civil rights movement. It produced the end of school segregation and a diminution of employment discrimination. Southern states began spending more money on education. A "northernization" of southern society, brought about by white in-migration and an influx of northern manufacturing firms, probably helped to create demand for better schools. Better education meant better pay, better jobs, more professional employment for blacks and whites, and—perhaps most important—greatly improved chances of getting jobs in the first place. The weakening of employment discrimination further

helped the newly educated blacks to get jobs. The civil-rights-based gains in education and employment led, in turn, to higher per-capita employment in the South. Urbanization also contributed to this last development, since cities have more jobs for women and teenagers.

A third post-1950 development was rapid industrialization—manufacturing growth—in the South. The industrialization was spurred by the South's strong markets (low ratios of "home grown" supply to demand), weak unions, low wages, and mild winters. New manufacturing meant more than just better-paying jobs to replace the ones lost in agriculture. It also meant a more urban society, one with more urban type occupations, including retail and service jobs required to accommodate the spending of the new factory workers. And the new factories meant more jobs for women, especially in the textile and apparel industries but also in retail and service jobs that accompanied urbanization. More jobs for women, in turn, meant higher labor force participation rates and higher employment ratios.

A final post-1950 development in the South was an upsurge in wage levels. Perhaps the main factor behind this upsurge was the loss of the South's huge labor reserve in agriculture. A complementary factor was the increased demand for labor in manufacturing, retailing, services, and other nonagricultural sectors. Also, in-migrants from the North brought with them wage-salary expectations conditioned by their former experiences. And some of the new manufacturing plants were covered by industry-wide collective bargaining agreements that required the same wages (northern wages) in all parts of the country.

The Nonsouthern Regions

Developments in the nonsouthern regions also contributed to income convergence. One such development was the technological revolution in agriculture. This revolution was not confined to the South; it was nationwide. And in the nonsouthern regions it differed enough from what happened in the South to be regarded as a somewhat different phenomenon. For example, the mechanical cotton picker—the preeminent technological advance in the South—was not a factor outside the South. Similarly, the nonsouthern revolution did not involve sharecroppers: there were none. The non-South did have tenant farmers; but they were a wealthier breed than their southern counterparts, and they were cash renters. (Southern tenants often paid with crop shares.) In the North, the technological revolution was characterized by extensive mechanization; extensive use of fertilizers, herbicides, and pesticides; and improved—more bountiful, more disease-resistant—crops. Farm consolidation and corporate farming were important aspects of the technological revolution. They helped spread the cost of mechanization over more acreage, making new equipment affordable.

Under the impetus of the new technology, the nation's farm population fell sharply and steadily. The nonsouthern region that benefited most was the Plains. It was the most agricultural region outside the South and, except for the Rockies, had the lowest 1950 per-capita income outside the South. It thus had the most to gain—and did gain the most—from the nonsouthern shift from agriculture to a more urban mix of industries and occupations.

The regions that benefited least from the agricultural revolution were the Manufacturing Belt and the Rockies. As the nation's most industrialized region, the Manufacturing Belt had comparatively little agriculture; there was less potential for farm-to-urban population shifts. In the West's Rocky Mountain subregion, the farm-to-urban population shift was restrained by a qualitative difference in agriculture. Large parts of the Rockies are mountainous and arid, poorly suited for crop farming. In these areas the predominant form of agriculture is ranching, the raising of sheep and cattle. Combines, fertilizers, pesticides, and herbicides are of no value in ranching, although some gains did occur in the field of animal genetics. (Also, the Jeep tended to replace the horse as the steed of the cowboy.) Apart from the small size of the farm-to-urban shifts in the Rockies, the shifts that did occur had relatively minor income effects. Rocky Mountain farm incomes were relatively high—70 percent of nonfarm incomes in 1967, compared to 52 percent for the United States as a whole and to less than 40 percent for some southern states.[35]

Another nonsouthern development affected only the West. The transport cost element of the price of manufactured goods shrank, causing relative prices to fall. In 1950 the West was the nation's least industrialized region. Only 4 percent of the West's population worked in manufacturing in 1950, compared to 5 percent of the Plains, 6 percent of the South's, and 14 percent of the Manufacturing Belt's.[36] A large share of the West's manufactured goods had to be imported from the Manufacturing Belt—the northeastern United States. These goods had to be shipped extremely long distances. So they bore long-distance transportation costs. These costs raised the price of goods sold in the West, thereby raising the West's cost of living. We see this same phenomenon—a high cost of living in a remote region—in Alaska today: the 1987 per-capita income of Alaska was exceeded by that of only three other states (Connecticut, New Jersey, and Massachusetts). The West's wage-salary structure adjusted to the cost of living, giving the West its high 1950 per-capita income—high despite the region's nonindustrial status.

Two things happened that shrank the transport cost element in western prices, lowering the West's cost of living. First, the West became more self-sufficient in manufactured goods. Throughout the post-1950 era, the West had higher manufacturing growth rates than any other region.[37] By producing more of its own goods, the West eliminated the long-distance shipping costs embodied in the prices of many goods. Second, transportation improvements brought significant reductions in the cost of shipping the remaining goods to the West. The 41,000 mile interstate highway system, begun in 1956, collaborated with larger, longer tractor-trailers to cut highway shipping costs. The railroads cut costs by converting from steam to diesel locomotives, adopting bigger freight cars, laying continuous rail, automating their freight yards, eliminating union featherbedding, and eliminating passenger trains (which caused freight trains to be sidetracked and delayed).

The lower cost of living that these changes gave the West had an effect. Relative wages and salaries fell, adjusting to the relatively lower cost of living. Average hourly wages in manufacturing for the West went from 114 percent of the national average in 1950 to 103 percent in 1987. This change was not part of any general wage decline in the high-income regions: wages in the industrially

self-sufficient Manufacturing Belt went *up*—from 104 percent to 105 percent of the national average. The result of the West's relative wage-salary decline was a decline in relative per-capita income. The income decline nearly matched the manufacturing wage decline. The West's per-capita income went from 114 percent of the national average in 1950 to 105 percent in 1987.

Migration developments probably had additional effects on western income, but the effects were mixed. Three migration flows are especially relevant. First, working-age persons migrated to the West from the industrially stagnant Manufacturing Belt and the agriculturally declining Plains states.[38] The effect of this flow was to raise the West's working-age population percentage, thereby raising per-capita employment and per-capita income. Second, affluent retirees migrated to the West in search of amenities. Colorado, New Mexico, Arizona, California, and Oregon proved especially attractive. Because they had above-average incomes, the retirees tended to cause per-capita income gains. Offsetting these two flows—flows of displaced workers and affluent retirees—was the inflow of Hispanics from South and Central America. The Hispanic immigrants, mostly illegal aliens, were poor; their effect was to reduce the West's per-capita income. These negative effects were greatest in the Mexican border states of New Mexico, Arizona, and California. If we can judge from the fact that the West's 1950–87 decline in relative income was slightly less than its decline in relative wages, the net effect of the three migration flows was positive.

Income convergence was helped by one other nonsouthern development besides the agricultural revolution, western transportation cost reductions, and interregional migration: educational gains in the Plains. The educational gap between the Plains and the other nonsouthern regions virtually disappeared. During 1950–80 median years of schooling went from 9.0 to 12.5 in the Plains, from 9.6 to 12.5 in the Manufacturing Belt, and from 11.3 to 12.7 in the West. The above-average gains in the Plains resulted from (*a*) above-average declines in farm population—farm children have less schooling, (*b*) school district consolidation and the concomitant disappearance of the one-room Little Red Schoolhouse, and (*c*) state equalization aid to school districts. Naturally, the relatively fast rate of educational improvement in the Plains helped the region achieve its relatively fast income growth.

The Overall United States

Beyond the southern and nonsouthern developments lie some general developments that operated more or less nationwide. The first of these, retiree migration, has already been mentioned in connection with migration flows into the West. During the study period the nation's population grew wealthier and more mobile. Retirees in general are a low-income group. They don't work, and their retirement benefits are often limited to social security. But the retirees who migrate tend to be affluent; they constitute a high-income group.[39] Certain amenity-rich states like Florida, Arkansas, Arizona, and California have attracted the lion's share of these affluent retirees. The amenities in question are lakes, mountains, seacoasts, and mild winters. In theory, the states that have attracted

retirees should have thereby raised their per-capita incomes.

A second general development is growth in the population percentage of widowed and divorced women. Divorce rates have been rising steadily for more than a century. Present divorce rates are twice those immediately following World War II and three times those of the 1920s and 1930s.[40] And the aging of the general population has inflated the population percentage of widows. Both widowed and divorced women tend to have low income.[41] Nineteen percent of widowed women and 24 percent of divorced women age 45 and over lived in poverty in 1989. And the risk of poverty was increased by older age and minority status; 42 percent of divorced black women age 62 and older had poverty-level incomes in 1989. High widowed-divorced percentages therefore affect income adversely.[42] Unevenness among states in their 1950–87 rates of change in the widowed-divorced percentage should contribute to variation in income growth rates.

A third general development is temporary economic fluctuations at the state and subregional levels. Coughlin and Mandelbaum identified three types of fluctuations; all three took place in the 1980s.[43] In 1987 New England was in the midst of a short-lived economic boom; the Mid-Atlantic states shared some of this good fortune. Defense contracts and growth in high-tech industries looked like the chief causes of the boom. Elsewhere, mainly in the South and West, oil-producing and coal-producing states fell on hard times. The main problem was oil price declines after 1981, when the OPEC cartel began losing its grip on the artificially high prices it had set. Depletion of domestic petroleum reserves compounded the problem. Since coal is a substitute for oil, the coal-producing states suffered with the oil-producing states when energy prices declined. The third temporary fluctuation of the 1980s was the farm crisis. Declines in farm exports caused agricultural depression, reducing income in farm states.

NOTES

1. Eli Heckscher, "The Effect of Foreign Trade on the Distribution of Income," *Ekonomisk Tidskrift* 21 (1919), 497–512; and Bertil Ohlin, *Interregional and International Trade* (Cambridge, Mass.: Harvard University Press, 1933).

2. The migration rates are from the *County and City Data Books* (Washington, D.C.: Bureau of the Census) for 1967, 1977, and 1988, and from John F. Long, *Population Deconcentration in the United States* (Washington, D.C.: Bureau of the Census, 1981), table 7.

3. William H. Crown and Stephen F. Fournier, *Migration and Regional Development* (Washington, D.C.: Economic Development Administration, 1987), table 14.

4. Edward L. Ullman, "Amenities as a Factor in Regional Growth," *Geographical Review* 44 (1954), 119–32; Lee D. Olvey, "Regional Growth and Inter-Regional Migration—Their Pattern of Interaction," *Review of Regional Studies* 2 (1972), 139–63; Bernard M. Bass and Ralph Alexander, "Climate, Economy, and the Differential Migration of White and Nonwhite Workers," *Journal of Applied Psychology* 56 (1972), 518–21; Larry M. Svarte, "Environmental Preference Migration: A Review," *Geographical Review* 66 (1976), 314–30; Rjindar K. Koshal, "Urban Growth and

Environment," *Growth and Change* 7 (1976), 34–38; Larry H. Long and Kristin A. Hansen, *Reasons for Interstate Migration: Jobs, Retirement, Climate, and Other Influences* (Washington, D.C.: Bureau of the Census, 1979); Lynn E. Browne, "The Shifting Pattern of Interregional Migration," *New England Economic Review* (November–December 1979), 17–32; Philip E. Graves, "Migration and Climate," *Journal of Regional Science* 20 (1980), 227–37; Leonard F. Wheat, John M. Wardwell, and Lee Faulkner, *The 1970s Migration Turnaround in Rural, Nonadjacent Counties* (Washington, D.C.: Economic Development Administration, 1984); and Brian J. Cushing, "Location-Specific Amenities, Topography, and Population Migration," *Annals of Regional Science* 21 (1987), 74–85.

5. Lynn E. Brown, "Narrowing Regional Income Differentials," *New England Economic Review* (September–October 1980), 35–54, and (November–December 1980), 40–57.

6. Glenn E. McLaughlin and Stefan Robock, *Why Industry Moves South* (Washington, D.C.: National Planning Association, 1949); Wilbur R. Thompson and John M. Mattila, *An Econometric Model of Postwar State Industrial Development* (Detroit: Wayne State University Press, 1959); Harvey S. Perloff, Edgar S. Dunn, Jr., Eric E. Lampard, and Richard F. Muth, *Regions, Resources, and Economic Growth* (Baltimore: Johns Hopkins Press, 1960); and Leonard F. Wheat, *Regional Growth and Industrial Location* (Lexington, Mass: D. C. Heath, 1973).

7. Wheat, *Regional Growth*; Wheat, "The Determinants of 1963–77 Regional Manufacturing Growth: Why the South and West Grow," *Journal of Regional Science* 26 (1986), 643–44; and William H. Crown, *Manufacturing Employment Growth in Urbanized Nonmetropolitan Counties* (Washington, D.C.: Economic Development Administration, 1989).

8. Crown, *Manufacturing Growth*, table 2.1.

9. McLaughlin and Robock, *Industry*; Thompson and Mattila, *Econometric Model*; Melvin L. Greenhut, "An Empirical Model and a Survey: New Plant Locations in Florida," *Review of Economics and Statistics* 41 (1959), 433–38; Benjamin Chinitz and Raymond T. Vernon, "Changing Forces in Industrial Location," *Harvard Business Review* 38 (1960), 126–36; Robert M. Lichtenberg, *One-Tenth of a Nation* (Cambridge: Harvard University Press, 1960); Perloff et al., *Regions*; Fantus Company, *The Appalachian Location Research Studies Program: Summary Report and Recommendations* (Washington, D.C.: Appalachian Regional Commission, 1966); Wheat, *Regional Growth*; Robert W. Schmenner, *The Manufacturing Location Decision: Evidence from Cincinnati and New England* (Washington, D.C.: Economic Development Administration, 1978); Howard A. Stafford, *Principles of Industrial Facility Location* (Atlanta: Conway Publications, 1979); James O. Wheeler, "Effects of Geographical Scale on Location Decisions in Manufacturing: The Atlanta Example," *Economic Geography* 57 (1981), 134–45; Howard A. Stafford, "Environmental Protection and Industrial Location," *Annals of the Association of American Geographers* 75 (1985), 227–49; R. D. Norton, "Industrial Policy and American Renewal," *Journal of Economic Literature* 24 (1986), 1–40; Wheat, "Manufacturing Growth"; and Neal E. Duffy, "The Determinants of State Manufacturing Growth Rates: A Two-Digit-Level Analysis," *Journal of Regional Science* 34 (1994), 137–62.

10. Chinitz and Vernon, "Changing Forces"; Wheat, *Regional Growth*, and "Manufacturing Growth"; Norton, "Industrial Policy"; Perloff et al., *Regions*; Crown, *Manufacturing Growth*; and John E. Connaughton and Ronald A. Madsen, "The Changing

Regional Structure of the U.S. Economy," *Growth and Change* 21 (1990), 48–60.

11. Wheat, *Regional Growth*, and "Manufacturing Growth."

12. Mancur Olson, *The Rise and Decline of Nations: Economic Growth, Stagflation, and Social Rigidities* (New Haven: Yale University Press, 1982), 78, 211. In an earlier version of this theory, Olson used the term *institutional arthritis*. See Mancur Olson, "The Causes and Quality of Southern Growth," in E. Blaine Liner and Lawrence K. Lynch (eds.), *The Economics of Southern Growth* (Research Triangle Park, N.C.: Southern Growth Policies Board, 1977).

13. McLaughlin and Robock, *Why Industry Moves South*; Thompson and Mattila, *Econometric Model*; Melvin L. Greenhut, "An Explanation of Industrial Development in Underdeveloped Areas of the United States," *Land Economics* 36 (1960), 371–79; Victor R. Fuchs, *Changes in the Location of Manufacturing Since 1929* (New Haven: Yale University Press, 1962); Wheat, *Regional Growth*; Schmenner, *Manufacturing Location Decision*; Richard E. Lonsdale and H. L. Seyler, *Nonmetropolitan Industrialization* (New York: John Wiley & Sons, 1979); Stafford, *Principles of Industrial Facility Location*; Schmenner, *Making Business Location Decisions* (Englewood Cliffs, N.J.: Prentice-Hall, 1982); Leonard F. Wheat, "The Determinants of 1963–77 Regional Manufacturing Growth," *Journal of Regional Science* 26 (1986), 635–59; and James H. Rubenstein, *The Changing U.S. Auto Industry: A Geographical Analysis* (London: Routledge, 1992).

14. Wheat, *Regional Growth*.

15. Olson, *Rise and Decline*, 53.

16. Ibid., 53–55.

17. Ibid., 53.

18. Ibid., 63.

19. Ibid.

20. John Kenneth Galbraith, *American Capitalism: The Concept of Countervailing Power* (Boston: Houghton Mifflin, 1952), 94.

21. Ibid., 94–95.

22. Olson, *Rise and Decline*, 63.

23. Ibid., 75, 105, 72.

24. David B. Truman, *The Governmental Process: Political Interests and Public Opinion* (New York: Alfred A. Knopf, 1951); Galbraith, *Capitalism*.

25. Howard A. Stafford, "Environmental Protection and Industrial Location," *Annals of the Association of American Geographers* 75 (1985), 227–49.

26. Leonard F. Wheat, "The Trouble with Sclerosis: What Causes Regional Growth Differences" (paper presented at the Regional Science Association Meetings, Boston, 1990).

27. Crown, *Manufacturing Growth*; Wheat, *Regional Growth*, and "Manufacturing Growth"; and Connaughton and Madsen, "Changing Regional Structure."

28. See note 9.

29. Wheat, *Regional Growth*, for 1947–63, and Wheat, "Manufacturing Growth," for 1963–77.

30. Fuchs, *Changes in Location*.

31. Olson, *Rise and Decline*, p. 258, n. 43.

32. James G. Maddox, with E. E. Liebhafsky, Vivian W. Henderson, and Herbert M. Hamlin, *The Advancing South: Manpower Prospects and Problems* (New York: Twentieth Century Fund, 1967), 8.

33. Gavin Wright, *Old South, New South: Revolutions in the Southern Economy Since the Civil War* (New York: Basic Books, 1986).

34. Ibid., 79.

35. Browne, "Differentials," 41, 43.

36. Wheat, *Regional Growth.*

37. Perloff et al., *Regions*, 1960; Wheat, *Regional Growth*, and "Manufacturing Growth"; Crown, *Manufacturing Growth*; and Robert W. Crandall, *Manufacturing on the Move* (Washington, D.C.: Brookings Institution, 1993), 1967–91 percentage changes computed from table 1-7.

38. Crown, *Migration*, table 14.

39. William H. Crown, "State Economic Implications of Elderly Interstate Migration," *The Gerontologist* 28 (1988), 533–39; B. Richards, "Old Money: An Influx of Retirees Pumps New Vitality into Distressed Towns," *Wall Street Journal*, August 5, 1988; W. Hass, III and W. Serow, *The Influence of Retirement In-migration on Local Development*; Final Report to the Appalachian Regional Commission (Asheville, N.C.: University of North Carolina), 1990.

40. William Goode, *The Family* (Englewood Cliffs, N.J.: Prentice-Hall, 1982) and Roy Marshall, *The State of Families* (Milwaukee: Family Service America, 1991).

41. William H. Crown, Phyllis H. Mutschler, James H. Schulz, and Rebecca Loew, *The Economic Status of Divorced Older Women* (Waltham, Mass: Policy Center on Aging, Brandeis University, 1993).

42. Leonard F. Wheat, *Labor Force Participation and Unemployment in American Counties* (Washington, D.C.: Economic Development Administration, 1982), Appendix A, "Determinants of Median Family Income"; and Kent P. Schwirian, *Determinants of County and City Median Family Income and Poverty Rates* (Washington, D.C.: Economic Development Administration, 1991).

43. Cletus C. Coughlin and Thomas B. Mandelbaum, "Why Have State Per Capita Incomes Diverged Recently?" *Federal Reserve Bank of St. Louis Review* 70, no. 5 (1988), 31–35.

4

HYPOTHESES AND PRELIMINARY EVIDENCE

In this study we test ten general hypotheses and a larger number of subhypotheses that try to explain state per-capita income growth. The general hypotheses attribute changes in per-capita income to *changes* in (1) farm–urban population mix, (2) educational attainment, (3) worker-nonworker population mix, (4) the black population percentage, (5) the population percentages of other low-income and high-income groups, such as widowed and divorced women, (6) hourly wage rates, (7) the proportionate amount of transportation costs embodied in the price of manufactured goods sold in the West, (8) industry mix, (9) energy prices, and (10) short-run economic conditions affecting only certain regions and states. The subhypotheses look at specific facets of the general hypotheses—for example, the college education facet of educational attainment. We also test some interaction hypotheses that assign stronger effects to some of the hypotheses in either the former slave (southern) or nonslave (nonsouthern) states.

RURAL-URBAN MIX

Urban areas, and particularly metropolitan cities, have generally higher per-capita incomes than rural areas. The relationship between urbanization and per-capita income is easily discerned. Consider these facts:

- In 1950, 5 of the 7 states with the highest urban population percentages were among the seven with the highest per-capita incomes.
- In 1950, 6 of the 9 states with the lowest urban percentages were among the nine with the lowest incomes.
- In 1980, (*a*) 3 of the 4 states with the highest urban percentages were among the four with the highest incomes, and (*b*) 8 of the 12 states with the highest urban percentages were among the eleven with the highest incomes.

- In 1980, all 8 states with the lowest urban percentages were among the thirteen with the lowest incomes.
- Wheat found that two urbanization variables—the log of population and the urban population percentage—had respective simple correlations of +.55 and +.52 with 1969 county median family income (2,706 counties).[1]
- Schwirian found that two urbanization variables—a metropolitan county dummy and the urban population percentage—had correlations of +.50 with 1979 county median family income (3,138 counties).[2]
- Table 2.1 shows that counties in the lowest of nine 1969 median income intervals had an average population of 10,544, compared to 472,294 for counties in the highest income interval. There was a near-perfect (one exception) rank relationship between income and average population.
- Table 4.1 shows that counties in the lowest of ten urban population percentage intervals had an average median family income of $6,388, compared to $10,261 for counties in the highest urban interval. Except in the third-lowest interval, each increase in urban percentage is associated with an increase in income. In the exception, the urban percentage increases, but the other three measures of urbanization—population, population density, and farm population percentage—show decreases in urbanization. These decreases explain the decrease in income.

Although the preliminary evidence listed here has nothing held constant, we shall find that the relationship between urbanization and income holds up solidly in multiple regression testing.

The reason for the preliminary findings is easy to grasp. Personal earnings are lowest among farmers and farm workers, somewhat higher among rural nonfarm residents (some of whom are merchants or factory workers), higher still among urban nonmetropolitan residents (who have higher proportions of professional and technical workers and higher female labor force participation rates), and highest of all among metropolitan residents. A 1966 study thus found: "Non-farm rural income is below that of urban dwellers, but is substantially higher than that of farm people."[3] Specifically, the study found that 1959 median income for farm families was only 57 percent of U.S. median income; rural nonfarm median income was 84 percent of the national median.

In 1950, the very lowest earnings were associated with the South's farm sector, which was dominated by sharecroppers and tenant farmers. These farmers included both blacks (the descendants of slaves) and poor whites. The black and poor white sharecroppers and tenant farmers often had little or no cash income. Poverty based on sharecropping and tenant farming was rampant in the South and was certainly the leading factor behind the South's low per-capita income.

At the other end of the rural-urban spectrum, metropolitan cities and counties—roughly the three highest intervals in table 4.1—have the highest per-capita incomes. Three things, one of which is two-faceted, explain this:

Table 4.1
Average Values of Selected Variables for Counties in Ten 1970 Urban Population Percentage Intervals

Urban Population % Interval	Number of Counties	Median Family Income	Population	Population per Sq. Mile	Population % Farm	Median Years School	Labor Force % Female	E/P[a]
90-100%	99	$10,261	693,033	3,910	0.2%	11.7	39.0%	40.1%
80-89.9	77	9,889	241,336	447	2.0	11.9	37.5	38.7
70-79.9	104	8,888	114,443	194	4.9	11.6	37.0	37.3
60-69.9	172	8,254	60,766	101	8.7	11.5	36.5	37.1
50-59.9	252	7,893	43,332	69	10.6	11.2	36.1	36.4
40-49.9	331	7,679	34,687	58	12.9	11.0	36.0	36.0
30-39.9	362	7,332	27,874	48	15.0	10.7	35.8	35.5
20-29.9	301	7,018	24,266	41	17.4	10.4	35.2	34.6
10-19.9	142	7,179	27,111	48	15.0	10.4	34.7	34.2
0-9.9	866	6,388	8,685	17	24.6	10.4	32.6	33.9

[a]The ratio of employment (E) to population (P), expressed here as a percentage. That is, it is the percentage of the population that is employed (hence earning wage-salary income).

Note: Includes all U.S. counties and county equivalents except (*a*) 275 counties where the labor force was less than 98 percent civilian (military base counties), (*b*) 121 counties where 10 percent or more of the population was college students, (*c*) 22 counties where 20 percent or more of the population was sixty-five or older *and* the net migration rate was +10 percent or more (retirement counties), and (*d*) 28 nonmilitary counties with male population percentages of 53 percent or more (counties with penal institutions, men's colleges, military schools, a boy's ranch, or hardly any people).

Source: Leonard F. Wheat, *Labor Force Participation and Unemployment in American Counties* (Washington, D.C.: Economic Development Administration, 1982), table 3-4 (median income and E/P are from unpublished tabulations).

1a. Metropolitan areas have better occupational mixes. They have proportionately more professional, managerial, and technical workers; more skilled tradesmen such as plumbers, mechanics, and machinists; and more equipment operators.

1b. As a corollary of better occupational mix, metropolitan areas have better-educated populations: most professional and managerial workers and many technical workers are college educated. Table 4.1 thus shows that each increase in urban population percentage (bottom of table to top) brings an increase in median years of schooling, except that the two lowest intervals are essentially equal in schooling and the highest interval shows a slight decline. (The educational decline in the highest urban interval results from central city ghetto populations: the black population percentage goes from 6.4 percent in the second-highest interval to 13.4 percent in the highest interval.)[4]

2. Metropolitan areas have more jobs for women and teenagers. The wider availability of female-type jobs puts proportionately more women in the labor force and thereby raises employment/population (E/P) ratios. Table 4.1 thus shows that, without exception, each increase in urbanization brings an increase in both the percentage of the labor force that consists of women and the E/P ratio. The E/P increases are exceptionally large in the two highest urban intervals.

3. Metropolitan areas have higher wage-salary structures, epitomized by the extremely high wage-salary structure of New York City. Higher wages are needed to attract and hold workers in the face of higher living costs. High housing costs and business rents are important elements of the high metropolitan cost of living. Competition for relatively close-in housing drives up housing costs—and forces people who pay less for housing (remote from downtown) to pay more for transportation. Unions, generally stronger in metropolitan areas, raise construction costs and thereby add further to housing and business rental costs. (Unions also affect income directly.) Higher taxes also add to the metropolitan cost of living. The higher taxes result from higher levels of governmental services—fire, police, water, sewer, library, and welfare outlays, for example. And goods tend to cost more because of high rentals, high business taxes, and union labor as elements of business costs.

Between 1950 and 1987 the United States had a prodigious rural-to-urban shift in population. America's farm population declined from 15 percent of total population in 1950 to 2 percent in 1987.[5] The metropolitan population percentage increased from 56 percent in 1950 to 77 percent in 1987.[6] The urban population percentage went from 64 percent in 1950 to 73 percent in 1987.[7]

The text table that follows compares the 1950 and 1980 urban percentages for two census regions (Northeast and West) and two subregions in both the South and North Central regions.[8] (The 1987 urban percentages are not available at the state level.) In many ways the South shows important differences between the Deep South (North Carolina, South Carolina, Georgia, Tennessee, Alabama, Mississippi, Arkansas, and Louisiana) and the South's peripheral states (Delaware, Maryland, Virginia, West Virginia, Kentucky, Oklahoma, Texas, and Florida). The Peripheral South is transitional and compromises between southern and northern or western characteristics. The values shown below (and in subsequent text tables) for the Deep South and Peripheral South are unweighted averages of the state values for the states included. The North Central region is also divided into subregions—the regular census divisions. This North Central breakdown is important, because the East North Central division is in the highly urbanized Manufacturing Belt, whereas the West North Central division is relatively agricultural. The 1950 and 1980 regional urban population percentages are as follows:

Region	1950	1980
Deep South	39.9%	56.6%
Peripheral South	53.7	65.7
Northeast	79.5	79.2
Great Lakes (East North Central)	69.7	73.3
Plains (West North Central)	52.0	63.9
West	69.8	83.9

The table shows that the urban gains were far larger in the Deep South than elsewhere. The urban percentage increased 42 percent in the Deep South, 23 percent in the Plains, 22 percent in the Peripheral South, 20 percent in the West, 5 percent in the Great Lakes, and not at all in the Northeast (0.4 percent decline).

A technological revolution in agriculture was largely responsible for the relative shift in urban population. Agriculture became more mechanized, more chemically oriented (fertilizers, herbicides, pesticides), and better supported by genetically engineered crops. A specific aspect of mechanization was farm consolidation. Consolidation produced bigger farms, many of which were commercial farms that replaced family farms. The larger farms, by spreading the cost of equipment over more acreage, facilitated a higher degree of mechanization. The larger farms were more profitable. They brought higher incomes to the farmers who remained. The upshot of these developments was a more efficient system of agriculture that could produce more despite using far less labor. Farmers and farm workers were forced off their farms and into cities.

Agriculture's technological revolution had the most effect in the South. Southern agriculture was the most labor intensive, because southern labor—mostly sharecroppers and low-paid hourly workers—was so cheap. The mechanical cotton picker, chemical control of weeds (replacing hoeing), and similar developments sent millions of southerners off the farms. Many went to the North; others went to southern cities. Either way, the result was increases in the South's urban and metropolitan population percentages. And that marked a shift in employment away from poverty-level agricultural work to generally better-paying urban jobs.

Outside the South the farm–urban earnings disparity was not nearly as severe. And the decline in farm population was less precipitous. Hence the nonsouthern effects of the rural-to-urban population shift should not be as severe. But they should definitely be noticeable.

These rural-urban population developments support the subordinate hypotheses that per-capita income increases should be high where

(1) decreases in farm population are big,

(2) decreases in tenant farm population (which counts sharecroppers) are big,

(3) relative increases in rural nonfarm population are low,

(4) relative increases in urban population are high,

(5) relative increases in metropolitan population are high, and

(6) increases in average farm acreage are high.

In addition to these six relationships, there may be more complicated ones involving interaction between acreage increases and farm population decreases. The acreage increases that resulted from farm consolidation raised the income of farmers. Per-capita income increases should be especially large in states where large reductions in farm population are associated with large increases in acreage. The interaction between farm population and acreage can be represented mathematically by multiplying farm population change by acreage change. The effect of multiplying farm population change (a negative value) by acreage change (positive value) is to make the biggest farm population decreases (negative values) *relatively* even bigger (more negative) when those farm decreases are associated with relatively large acreage increases. If farm change and acreage change interact, per-capita income increases should be high where

(7) a high (extremely negative) product results from multiplying farm population change by acreage change and

(8) a high (extremely negative) product results from multiplying tenant farm population change by acreage change.

EDUCATION

It is well known that high educational attainment is associated with high income: better-educated people, especially college graduates, tend to earn more. Education therefore should affect per-capita income. Here is some evidence:

- In 1950, 5 of the 8 states with the highest median years of schooling (persons 25 and older) were among the ten with the highest per-capita incomes.
- In 1950, 10 of the 11 states with the lowest median years of schooling accounted for all ten with the lowest per-capita income.
- In 1980, 7 of the 10 states with the highest percentages of college graduates (persons 25 and older) were among the ten with the highest per-capita income.
- In 1980, 8 of the 9 states with the lowest percentages of college graduates were among the nine with the lowest per-capita income.
- Wheat found that two education variables (along with a wage variable) had the highest correlations with 1969 county median family income: median years of schooling and the percentage of high school graduates both had correlations of +.73.[9]
- Schwirian found that two education variables had the highest and the third-highest correlations with 1979 county median family income: the percentage of adults with eight or fewer years of schooling registered -.68, and median years of schooling registered +.60.[10]

Table 4.2
Average Values of Selected Variables for Counties in Six 1970 Median Years of Schooling Intervals

Median Years of School Interval	Number of Counties	Median Family Income	Population % Urban	Population	Population per Sq. Mile	Population % Farm	E/P[a]	Population % Black
12.3+	103	$10,188	50.3%	232,988	321	8.7%	38.8%	1.3%
11.5-12.2	1,091	8,530	39.7	84,523	258	16.0	37.1	2.1
10.5-11.4	444	7,357	33.9	48,803	329	16.1	35.5	6.4
9.5-10.4	462	6,498	26.7	27,795	131	14.8	34.8	15.8
8.5-9.4	485	5,752	18.2	16,681	45	17.7	33.1	18.9
8.4 or less	121	4,588	13.9	15,702	33	19.2	18.6	16.4

[a]Ratio of employment (E) to population (P), expressed here as a percentage.

Note: Includes all U.S. counties and county equivalents except the 446 described in the note to table 4.1.

Source: Ibid., unpublished tabulations from the study.

- Table 2.1 shows that counties in the lowest of nine 1969 median income intervals had an average high school graduate percentage of 23.7, compared to 61.0 for counties in the highest interval. The high school graduate percentage increases in each successively higher income interval, without exception.
- Table 4.2 shows that counties in the lowest of six median-years-of-schooling intervals had an average median family income of $4,588, compared to $10,188 for counties in the highest educational interval. Each increase in schooling is associated with an increase in income, without exception.

Three things contribute to the relationship between education and income. First, well-educated persons are more likely to be working—earning income. They have more incentive to work: they have access to jobs with higher pay, interesting tasks, and attractive working conditions. Poorly educated persons, in contrast, can expect low pay, disagreeable or monotonous work, and possibly uncomfortable or hazardous working conditions. Beyond incentives, poorly educated persons—high school dropouts especially—often have trouble finding work. Some get discouraged and drop out of the labor force. The poorly educated persons who do find work usually have the types of jobs that subject them to periodic unemployment, especially during recessions. These considerations explain why, in table 4.2, a strong relationship between schooling and E/P (employment/population ratio) is evident. Reading the table from bottom to top, we see that each increase in median years of schooling brings an increase in E/P, without exception.

Second, and probably more important, education affects how much people earn when they do work. Someone without a high school diploma is likely to become a laborer, a gas station attendant, a waitress, or something of that sort. High school graduates have a much better chance of getting white-collar jobs; many become store managers or businessmen, and some get post-high-school technical educations that lead to moderately well-paying technical positions. College graduates have an excellent chance of entering professional or managerial positions with salaries in the upper-middle-income brackets; some become self-employed professionals—doctors, lawyers, architects—with high incomes. Since professional-technical-managerial employment is associated with urbanization, we can use urbanization as a proxy for occupational mix. Table 4.2 shows that each increase in schooling is associated with increases in both urban population percentage and population, without exception. This indirect evidence implies that well-educated persons tend to have professional-technical-managerial jobs, the kind of jobs found in urban areas.

Third, poorly educated persons are likely to be black, hence hampered by cultural barriers and racial discrimination in their efforts to find good jobs. Table 4.2 shows a powerful, if imperfect, relationship between median years of schooling and the black population percentage. The black percentages in the two lowest educational intervals provide a striking contrast with those in the two highest intervals.

Given education's effect on income, it is reasonable to suppose that rising educational attainment for a state should produce rising per-capita income. And educational attainment *has* been rising. More important, it has been rising in the South relative to the rest of the country.

The post-Civil War South put a low value on education, particularly for blacks but also for poor whites. Lemann observes that "whites kept the black school system in Mississippi inferior in part because they didn't want sharecroppers' children to have career options beyond sharecropping."[11] He cites a quotation from a 1939 white "creed of racial relations": "Negroes are necessary to the South and it is desirable that they should stay there and not migrate to the North."[12] Wright comments: "The economic structure that supported the low priority [for education in the South] was simply this: as a low-wage region in a high-wage country, the South had no expectation that it could capture the return on investment in its own people."[13] In other words, educated workers would be equipped for something better than sharecropping and would migrate in search of higher wages. Hence the sharecroppers and tenant farmers—and blacks generally—had little or no schooling.

Then four things happened. First, the sharecropper-tenant farmer system collapsed, and 5 million blacks left the South. Even without improvements in black educational attainment, the relative shift toward whites in the South's black-white population mix would have raised educational attainment levels in southern states. Northern levels, meanwhile, would have increased more slowly, because even in the North blacks get less education. And the sharecroppers who went north brought their lack of education with them; that pulled down on the northern state educational averages. Second, the civil rights movement produced school desegregation; southern blacks began to get better educations. Third, the federal

government and the states began providing more aid for education. Fourth, even while blacks were moving out of the South, whites were moving in. The whites had good educations—substantially above the southern level. The white in-migration helped raise the southern averages.

The result of these four developments was a remarkable evening out of educational attainment among the nation's regions. The following text table compares 1950 and 1980 median years of schooling (persons 25 and older) for the previously used regions and subregions:[14]

Region	1950	1980
Deep South	8.0%	12.2%
Peripheral South	9.0	12.4
Northeast	9.6	12.5
Great Lakes	9.6	12.4
Plains (West North Central)	9.0	12.5
West	11.3	12.7

These figures offer important revelations about both the South and the Plains. The Deep South—home of most of the nation's sharecroppers and blacks—was well behind the rest of the country in 1950; the Peripheral South also lagged. But by 1980 the Deep South had almost caught up, and the Peripheral South *had* caught up. The Plains also lagged in educational attainment in 1950, though not as much as the Deep South. Two things explain this lag. First, 1950 was still in the era of the Little Red Schoolhouse in the Plains; school district consolidation was just getting under way, and adequate-sized schools were not available to many rural children. Second, the Plains had proportionately more farmers than any other region outside the South. Farm families tended to place less value on education and were often concerned about having the children available for full-time help on the farm as early in their lives as possible; high school diplomas were not particularly useful in farming. By 1980 school district consolidation was completed, and rural children were commuting to large, graded, adequately equipped schools. Also, the region's farm population had fallen sharply.

The West's *relative* decline in educational attainment also merits comment. This decline seems to partly explain the West's parallel decline from national leadership in per-capita income. The West's high 1950 educational attainment is probably explained by two factors: an extremely low black population percentage and the lack of an offsetting high farm population percentage such as that of the Plains. Although not industrialized, the West was too arid and mountainous to be as agriculture-intensive as the Plains; the West's 1950 urban population percentage (69.8 percent) exceeded the nation's (64.0 percent), whereas the Plains's urban percentage (52.0 percent) was below that of any other region except the South (48.6 percent). By 1980, a large influx of blacks and Hispanics had given the West a higher minority population percentage—more like the minority percentages of the other nonsouthern regions. Since blacks and Hispanics have generally low educational attainment, the West's increase in minority

population helped bring the region's median years of schooling down (relatively) almost to the levels of the other regions.

These educational developments support the general hypothesis that high growth in educational attainment is associated with high growth in per-capita income. Three subordinate hypotheses (that relate to specific variables) are that per-capita income increases should be high where

(1) the increase in median years of school is high,
(2) the increase in the high school graduate percentage is high, and
(3) the increase in the college graduate percentage is high.

WORKER-NONWORKER MIX (EMPLOYMENT)

The third general hypothesis is that state per-capita income changes result from changes in the population mix of workers and nonworkers—that is, income earners and nonearners. This hypothesis divides population into three categories: employed persons, unemployed persons, and persons not in the labor force. (To be classified as unemployed, a person must be actively seeking work.) The employed and the unemployed persons constitute the labor force. The basic idea behind the hypothesis is that per-capita income should tend to be high where the proportion of wage-salary income earners in the population is high. Income should thus be high where per-capita employment (the employment/population ratio) and the labor force participation rate (the labor force/population ratio) are high and where the unemployment rate (unemployed persons/labor force) is low. Preliminary evidence suggests that the hypothesis is a strong one:

- In 1950, 6 of the 7 states with the highest per-capita employment (E/P) were among the seven with the highest per-capita income.
- In 1987, 7 of the 11 states with the highest per-capita employment were among the twelve with the highest per-capita income.
- In 1987, all of the 7 states with the lowest per-capita employment were among the eight with the lowest income.
- Wheat found that the labor force participation rate had an extremely significant correlation of +.63 with 1969 county median family income.[15]
- Schwirian found that the labor force participation rate had an extremely significant correlation of +.67 with 1979 county median family income.[16]
- Table 2.1 shows that counties in the lowest of nine 1969 median income intervals had an average labor force participation rate of 27.3 percent, compared to 41.5 percent for counties in the highest income interval. Over the nine intervals, each increase in income is associated with an increase in participation, without exception.
- Table 4.3 shows that counties in the lowest of seven employment/population ratio (E/P) intervals had an average median family income of $4,607, compared to $9,203

for counties in the highest E/P interval—a 100 percent difference in income. Each increase in E/P is accompanied by an increase in income, without exception.

Note in table 4.3 that, except among the smallest 500 or so counties (those in the three lowest E/P intervals), each increase in E/P is also associated with an increase in the labor force's female percentage. Among the larger, more urbanized counties, increases in E/P tend to result from increases in female labor force participation. Meanwhile, each increase in E/P is also associated with an increase in both the urban population percentage and median years of schooling, except that schooling deviates slightly in the highest E/P interval. These findings corroborate earlier ones that urbanization and education contribute to income by, among other means, raising per-capita employment. One other noteworthy finding is that, across the middle five E/P intervals, each increase in E/P is associated with a *decrease* in the black population percentage—evidence that blacks have difficulty finding work and that E/P is to some extent a proxy for the black percentage. (The black percentage rises in the highest E/P interval, because that is where most of the central cities are; see the population column. High female employment in the central cities offsets low black employment, permitting E/P to rise.)

E/P is the best single measure of the population mix of workers and nonworkers. The 1950 and 1987 E/P percentages for the regions and subregions used in previous comparisons follow.[17]

Region	1950 E/P	1987 E/P
Deep South	34.3%	59.0%
Peripheral South	35.3	60.2
Northeast	39.1	61.0
Great Lakes	39.1	61.0
Plains (West North Central)	38.1	64.8
West	36.1	62.8

Once again the South shows the least favorable 1950 circumstances. The Deep South is again off by itself at the low end of the scale. For both 1950 and 1987, the South's low E/P ratios were influenced by the fact that the South was (and is) the least urbanized region. On owner-occupied farms, women do work and they work hard, but they are not ordinarily counted as employed (wage earners). On sharecropper farms, the women worked as field hands. And in the cities, many women had jobs. Particularly in 1950, many urban women worked in traditional female jobs—as teachers, secretaries, nurses, clerks, waitresses, telephone operators, librarians, and the like. Today, proportionately more urban women have access to professional, managerial, and other better-paying jobs. Apart from jobs for women, cities also have more jobs for teenagers. The South's low E/P in 1950 reflected the proportionately small urban population, as partly offset by the stimulating effect of sharecropping on E/P.

Employment discrimination against blacks also contributed to the South's low

Table 4.3
Average Values of Selected Variables for Counties in Seven 1970 Employment/Population (E/P) Ratio Intervals

E/P Interval	Number of Counties	Median Family Income	Population % Urban	Population	Population % Farm	Median Years School	Population % Black	Labor Force % Female
41%+	193	$9,203	51.3%	203,730	8.6%	11.3	8.5%	39.4%
38–40.9	567	8,465	43.0	95,603	13.7	11.4	6.0	36.7
35–37.9	857	7,632	33.3	49,282	17.4	11.2	6.9	34.9
32–34.9	585	6,801	23.5	27,020	19.0	10.5	9.4	33.7
29–31.9	292	6,290	22.4	18,419	15.7	10.0	12.7	33.1
26–28.9	118	5,117	17.9	18,333	14.0	9.2	19.7	33.9
Below 26	94	4,607	11.5	17,061	15.3	8.6	12.3	31.6

Note: Includes all U.S. counties and county equivalents except the 446 described in the note to table 4.1.

Source: Ibid., unpublished tabulations from the study.

E/P; discrimination may have been the most important cause. Twenty-two percent of the South's 1950 population was nonwhite, and southern nonwhites constituted 66 percent of America's nonwhite population. Racial discrimination could therefore have had an appreciable effect on overall employment. In farm areas, discrimination was not a problem: the planters accepted blacks as a source of cheap labor.[18] But for urban blacks, the situation was entirely different. Employment discrimination was pervasive.[19] This problem was made worse by lack of education among blacks. They were effectively excluded from many occupational categories.

Between 1950 and 1987 the South almost caught up with the rest of the country in its E/P percentage. The Deep South was about 5 percentage points below the two Manufacturing Belt regions (Northeast and Great Lakes) in 1950 but was only 2 percentage points behind in 1987. What happened? Several things. First and most important, southern agriculture shrank drastically. As a result, the South became urbanized and metropolitanized more rapidly than the rest of the country; the South became more like the nonsouthern regions in its rural-urban mix. This change meant more jobs for women and teenagers. Second, previously discussed factors caused the South to catch up with the other regions in education. This development largely removed the South's educational obstacle to employment. Third, millions of blacks left the South, reducing the South's black population percentage while raising the black percentages of the other regions. To the extent that job discrimination and poor education were still barriers to black employment, the black population shift greatly reduced the barriers' effect in the South—and restrained the increases in E/P percentages in the regions the blacks moved to. Fourth, the civil rights movement, various court decisions, and new

civil rights laws curtailed employment discrimination against blacks. Those blacks who remained in the South found it easier to get jobs—all the more so because they were better educated in 1987 than in 1950. Fifth, the revolutionary increase in female labor force participation—beyond what urban growth can account for—enhanced the first factor: exceptionally rapid urbanization in the South. The huge size of the female labor force participation increase is illustrated by the 1950-87 increases in E/P shown in the text table.

The regions ranked as follows in percentage increase in E/P: West, 74 percent; Deep South, 72 percent; Peripheral South, 71 percent; Plains, 70 percent; Northeast, 56 percent; and Great Lakes, 56 percent. In other words, four regions benefited more or less equally at the expense of the Manufacturing Belt (Northeast plus Great Lakes). Three of the favored regions—Deep South, Peripheral South, and Plains—owed their E/P gains in large measure to their agricultural status: they had bigger farm-to-urban shifts than the Manufacturing Belt and (to a lesser extent) the West.

But what about the West? Since the West is less agricultural than the Plains, why did the West have a slightly larger (74 percent versus 70 percent) increase in E/P? The technological revolution in agriculture produced net out-migration, particularly among young adult males, from many Plains states in the 1950s. Since males have higher participation rates than females, this out-migration slowed the E/P increase in the Plains. The West, on the other hand, had the country's highest growth rates—in population, overall employment, and manufacturing employment. Abundant jobs in the West attracted the young male migrants and were a stimulus to employment, hence to E/P. Male migration's effect is visible in the fact that, in 1970, the West had the nation's highest male population percentage: 49.4 percent, compared to 48.7 percent for the United States.

Since the West declined relatively in per-capita income during 1950-87, the beneficial effects of its E/P gains were apparently offset by a cost-of-living reduction associated with transport cost reductions in the West and by the West's relative decline in educational attainment.

On the basis of this analysis, we can hypothesize that per-capita income gains will be high where

(1) E/P gains are high,
(2) labor force participation rate gains are high, and
(3) unemployment rate changes are low or negative.

BLACK POPULATION PERCENTAGE

Blacks tend to have low income—another well-known fact. We thus expect an association between the black population percentage and per-capita income. Preliminary evidence reveals no such association for high-income states. The apparent reason is that the high-income states are northern, and northern blacks gravitate to metropolitan areas, which have high income. The rural northern

states, whose lack of metropolitan areas limits them to moderate incomes, have the nation's lowest black percentages. But for low-income states and states in general, the expected relationship can be found:

- In 1950, the 7 states with the highest black percentages were among the ten with the lowest per-capita income; Mississippi, with the highest black percentage, had the lowest per-capita income.
- In 1980, 5 of the 9 states with the highest black percentages were among the six with the lowest per-capita income; Mississippi was again highest in black percentage and lowest in income.
- Wheat found a significant -.33 simple correlation between the black population percentage and 1969 median income (2,706 counties). The tendency of northern blacks to live in high-income areas— metropolitan areas—weakened the relationship. But when the log of county population density (urbanization) was held constant, the black percentage had a stronger partial *r* of -.47. When the high school graduate percentage was also controlled, the partial *r* fell to -.13: lack of education largely explains the effect of race on income.[20]
- Schwirian found correlations of -.23 and +.46 between the black percentage and, respectively, 1979 county median income and 1979 percentage of county families below the poverty level.[21] (The decline of the median income correlation from -.33 to -.23 between the 1970 and 1980 censuses reflects continued shifting of the black population from the rural South to high-income metropolitan counties in the Manufacturing Belt. The Mid-Atlantic black population percentage went from 10.6 to 11.9 during 1970-80, and the East North Central black percentage went from 9.6 to 10.9.)
- Table 2.1 shows that counties in the lowest of nine median income intervals had an average black population percentage of 24.2, compared to 3.3 in the third-highest interval and 4.3 in the highest. Except in the two highest income intervals, each increase in income is associated with a decrease in the black percentage.
- Table 4.4 shows that, if totally nonblack counties (lowest interval) are excluded from the analysis, counties in the lowest of six black population percentage intervals had an average median family income of $8,509, compared to $4,842 for counties in the highest interval (more than 50 percent black). Over the six intervals that include some black population, each increase in the black percentage is associated with a decrease in median income, without exception.

In the table 4.4 analysis, we are ignoring the totally nonblack counties because they are obviously in a class by themselves. The population and farm population percentage columns, viewed in conjunction with the fact that these counties have virtually no blacks, make it clear that these are nonsouthern agricultural counties. These counties don't fit the pattern of income change found in the table's other intervals. The reason: in these counties low income resulting from ruralness outweighs the income gains that might be expected from the black percentage's decline to essentially zero. As a related point, the educational gain that would

Table 4.4
Average Values of Selected Variables for Counties in Seven 1970 Black Population Percentage Intervals

Black Population % Interval	Number of Counties	Median Family Income	Median Years School	Population % Urban	Population	Population % Farm	E/P	Population % Under 18
50%+	97	$4,842	8.8	18.7%	17,133	14.4%	30.4%	40.9%
40–49.9	102	5,961	9.4	27.7	41,588	11.0	33.6	38.4
30–39.9	131	6,308	9.8	33.5	59,936	9.6	35.1	36.4
20–29.9	156	6,901	10.1	36.4	141,095	8.0	35.7	35.4
10–19.9	202	7,535	10.4	46.4	170,696	8.8	37.2	34.2
0.05-9.9	527	8,509	11.1	50.5	119,998	8.7	36.8	34.6
0–0.05	1,491	7,399	11.1	23.9	18,261	21.4	35.1	35.1

Note: Includes all U.S. counties and county equivalents except the 446 described in the note to table 4.1.

Source: Ibid., unpublished tabulations from the study.

ordinarily result from a decline in the black percentage (stub column) is neutralized by the educational loss that results from a big jump in the farm population percentage (farm column).

The reasons for the association between black percentage and income have been widely discussed and are generally understood. Blacks are hurt by poor education, employment discrimination, high fertility rates (high percentages of children, or nonworkers), and the so-called ghetto subculture. Poor education is especially important. Using county data, Wheat finds that the black percentage has correlations of +.61 with the population percentage of adults with less than five years of schooling and -.49 with the adult percentage of high school graduates.[22] And in table 4.4, each increase in the black percentage across the six nonzero black percentage intervals brings a decrease in median years of schooling: schooling drops from 11.1 years in counties that are less than 10 percent black to 8.8 years in counties that are more than 50 percent black. Ginzberg elaborates on the problem:

> Approximately one black youngster in four currently fails to graduate from high school. For many of them effective schooling really stops long before they leave or are pushed out, since a significant proportion become truants by the time they are 12 or 13 and many who remain in school pay little or no attention to their studies.[23]

Employment discrimination, the next handicap, is hard to document, but few people would deny its existence. High black fertility, another source of low per-capita employment, is reflected in the black percentage's correlations of +.29

with the percentage of population that falls in the 0–17 age bracket and -.26 with the population percentage in the 25–54 (prime working age) bracket.[24] Table 4.4 provides more detail: each increase in the black percentage over the six nonzero black population percentage intervals is accompanied by an increase in the population percentage under 18, except between the two lowest nonzero intervals.

For its part, the ghetto subculture has devastating effects. Lemann states that "the underclass culture in the ghettos is . . . the greatest barrier to progress by the black underclass."[25] This subculture has many facets, most of which tend to reduce black employment and earnings. One is a tendency toward welfare families headed by the mothers, who often have many children. The entire family consists of nonworkers. Resistance to education, especially among male children, interacts with language problems (black English) to make blacks relatively unemployable and to isolate them from jobs that pay well. This problem is made worse by gangs, which display antieducation attitudes, and by teenage pregnancy, which takes girls out of school. Above-average criminality among black males puts many in prison, or at least outside the labor force, when they might otherwise be working. In recent years, drug abuse has compounded the black employment problem.

These handicaps—poor education, job discrimination, high fertility, and the ghetto subculture—contribute to two other handicaps: low per-capita employment (low E/P) and low-paying jobs. The E/P column in table 4.4 shows that, except between the first two nonzero intervals, each increase in the black percentage is accompanied by a decrease in E/P. (E/P's increase from 36.8 to 37.2 between the first two nonzero intervals is explained by average population's increase from 119,998 to 170,696: more extensive metropolitanization offsets the effects of the higher black percentage.) In 1978 the E/P ratios for white young adults (20–24), both males and females, averaged 15 percentage points higher than those of blacks in the same age-sex cohorts.[26] As for job quality, one-third of all black workers are laborers or service workers; for whites the fraction is one-sixth.[27]

The black percentage's effect on income change takes on meaning when we examine the black population shifts among regions that occurred between 1950 and 1988. (No 1987 data are available.) The text table shows the 1950 and 1988 black percentages.[28]

Region	1950 % Black	1988 % Black
Deep South	30.5	25.7
Peripheral South	13.2	13.2
Northeast	5.1	11.0
Great Lakes	5.9	11.9
Plains (West North Central)	3.0	4.9
West	2.9	5.7

The Peripheral South did not change; black out-migration was neutralized by in-migration from the Deep South to such peripheral metropolitan areas as Richmond, Louisville, Oklahoma City, Dallas, Fort Worth, and Houston. But the

Deep South experienced a 15.7 percent reduction in its black percentage, enough to influence per-capita income profoundly. The Northeast and the Great Lakes had respective increases of 116 percent and 102 percent. These increases were capable of producing important income effects. The Plains and the West also had large increases—63 percent and 97 percent—although the percentage *point* increases were much smaller than in the other two nonsouthern regions; the income effects must have been less.

These interregional-transfer-of-poverty effects were strongest in the 1950s and early 1960s, when the black exodus from the South was at its height. The migration effects were aptly summarized in the 1967 report of the President's National Advisory Commission on Rural Poverty: "In the Deep South . . . a mass migration of Negroes, mainly to northern industrial centers, has helped reduce southern rural poverty at the expense of cities."[29]

OTHER LOW-INCOME AND HIGH-INCOME GROUPS

Blacks are not the only group whose proportionate size can influence per-capita income. We have put blacks in a separate category because they far outweigh the other groups in importance. But several other low-income and high-income groups have the potential to affect income and income change. Besides blacks, there are three low-income groups whose size might affect income: (1) widowed and divorced women, (2) Hispanics, and (3) Indians. At the other end of the income scale, affluent retirees who migrate to states with mild winters and other amenities could stimulate income growth.

The largest and potentially most important of these groups is widowed and divorced women. Prior research supports the idea that increases in the widowed-divorced percentage of total population will lower a state's relative income:

- Wheat found a significant simple correlation of -.33 between the widowed-divorced percentage and 1969 county median income. As with the black percentage, the relationship was partly obscured by the group's tendency to live in metropolitan areas, where income is high. Controlling the log of population density (urbanization) thus changed *r* from -.33 to -.49.[30] As table 2.2 shows, the widowed-divorced percentage was in both the best five-variable equation and the best ten-variable equation. And in both equations, female marital status was outranked in significance by only four other factors: farm-urban mix, education, labor force participation (roughly equivalent to E/P), and wage level.
- Schwirian found a slightly higher simple *r* of -.35 for the widowed-divorced percentage opposite 1979 county median income. When he looked at cities above 10,000, where divorce is less stigmatized and the widowed-divorced percentage therefore runs higher, he found a surprisingly high simple *r* of -.60 between that percentage and median income.[31]

The reasons for this income influence are not hard to grasp. Both widowed and divorced women tend to have low income.[32] Nineteen percent of widowed women

and 24 percent of divorced women age 45 and over lived in poverty in 1989. Moreover, the risk of poverty was increased by older age and minority status: 42 percent of divorced black women aged 62 and older had poverty level incomes in 1989. The soaring divorce rate and general aging of the population add to the potential importance of the widowed-divorced group.[33] Since divorce rates are higher in urban areas than in rural areas, the most urbanized regions—the Manufacturing Belt and the West—are likely to have had their high per-capita incomes pulled down toward the national average by divorce trends.

Less evidence exists about the income effects of the Hispanic and Indian populations. But these *are* low-income groups. Hispanics, like blacks, tend to be poorly educated and to be afflicted with the ghetto subculture, including its gang influences. Their lot is often made worse by an inability to speak English. The romanticized notion that America is a melting pot does not always agree with reality where Hispanics are concerned: those who are not native-born often have difficulty picking up the language. The resulting language barrier renders them unsuited for most better-paying jobs.

Indians also have a cultural problem, though perhaps not so much in the pejorative sense of a subculture. Many choose to live on reservations and to cling to their native culture. Living on the reservation limits their access to good jobs. And the culture, whatever its merits, is not one that equips a person for assimilation into modern society, for getting a good education, and ultimately for finding a job that pays well. Traditional crafts, for example, are seldom a good source of income. Despite these considerations, the Indian hypothesis seems weak. Whereas black and Hispanic migration is highly uneven geographically—Hispanics favor the Southwest, Florida, and large northeastern cities—Indian migration is more evenly distributed, not to mention its being negligible. Because of the tendency of Indians to live on reservations, the geographic distribution of this population is fairly stable—less likely to cause income change.

The last of our groups is at the opposite end of the income scale, the high end. This group is affluent retirees. Many of them move to states like Florida, Missouri, Arkansas, Arizona, and California—states with mild winters, seacoasts, lakes, and mountains. Whereas the Great Black Migration involved a geographical transfer of poverty, retiree migration involves a transfer of wealth. Retirees who migrate tend to be significantly above average in their incomes.[34] Whether retirees as a group are proportionately large enough to affect income significantly is more a matter of conjecture. But in the best-known and most extreme example, Florida, the 1980 percentage of persons over 65 in the population towered above the national average: Florida, 17.3 percent; United States, 11.3 percent.[35] (The second-highest over-65 percentage, 13.7 for Arkansas, was well behind Florida's.) Florida's 1950–80 percentage point *increase* in the over-65 percentage towered relatively even higher above the national average: Florida, 8.7 percentage points; United States, 3.2 percentage points. (Arkansas's increase was 5.9.) It is at least plausible, then, that retiree migration affected income.

The general hypothesis that per-capita income changes are related to changes in population percentages of low-income and high-income groups leads to several subordinate hypotheses (not counting the one for blacks). Per-capita income increases should be high where

(1) increases in the population percentage of widowed and divorced women are low,

(2) increases in the Hispanic population percentage are low

(3) increases in the Indian population percentage are low,

(4) the net migration rate for persons 65 and over is high, and

(5) amenity levels (retiree attractions) are high.

HOURLY WAGE RATES

Since the annual earnings of many workers depend largely on their hourly wage rates, it is reasonable to expect wage rate changes to be associated with income changes. There is, of course, no single wage rate: wage rates differ by occupation, industry, firm, rural-urban location, union-nonunion status, and so on. Yet there is a wage structure or hierarchy such that a given state's wage rates tend to be generally high or low across the board: a specific wage rate (e.g., manufacturing or operatives) can represent the hierarchy. Preliminary evidence supports the wage hypothesis:

- In 1950, 10 of the 12 states with the highest average hourly earnings in manufacturing (production workers) were among the twelve with the highest per-capita income. (No similar relationship can be found for 1987, however.)
- In 1950, 6 of the 7 states with the lowest manufacturing wages were among the twelve lowest in per-capita income.
- In 1987, 5 of the 10 states with the lowest manufacturing wage rates were among the twelve with the lowest income.
- Wheat found a +.73 correlation between median earnings for the "operatives" occupational group (a proxy for wages) and 1969 county median income. (A proxy was used because hourly wage data were not available for counties. The proxy is not wholly reliable, because it uses an element of income to predict income and because it is somewhat distorted by variation in hours worked per year.)
- Schwirian found a +.35 correlation between median earnings for operatives and 1979 median county income.

The observed relationships between wage *levels* and income *levels* suggest that increases in wage rates should cause increases in income. And hourly wage rates in the South have been catching up with wage rates in the rest of the country. In 1950 the Deep South's wage levels were far behind those in other regions. This is not surprising. Slavery gave the South extremely low wage rates: you can't earn much competing against unpaid slaves. (Slave owners sometimes rented out their slaves for factory and plantation work,[36] so there was the essence of a slave wage rate: but it was extremely low by northern wage standards.) Ever since the end of the Civil War, southern wages have been slowly—ever so slowly—creeping up toward the national average. But more than a century and a quarter

of change has not brought wage parity: the reserve labor in sharecropping saw to that. In 1950, or 85 years after the Civil War, the Deep South's wage rates still lagged far behind those of the rest of the United States. But by the late 1960s southern wages were climbing faster.[37] The Deep South's wage gap shrank.

The text table shows what happened. It compares 1950 average hourly earnings of manufacturing production workers with 1987 earnings; wages are expressed as percentages of the national average.[38]

Region	1950	1987
Deep South	80%	87%
Peripheral South	96	100
Northeast	98	98
Great Lakes	115	117
Plains (West North Central)	98	97
West	114	103

Only the Deep South and Peripheral South were more than 2 percentage points below the national average in 1950. The Deep South, which lagged the most in 1950, had moved up from 80 percent to 87 percent by 1987, a far bigger gain than found in any other region. The Peripheral South had reached the national average. Once more, the South's improvement can be traced in large measure to the demise of sharecropping. The South's large, undercompensated farm population gave the region a huge supply of reserve labor. When the sharecroppers left the farms, the immediate effect may have been to increase the availability of surplus labor. But many of these workers migrated north and west. And, over time, many others found jobs in the South's rapidly expanding manufacturing and service sectors. The gradual weakening of employment discrimination helped. So did improved education, which made the South's blacks and poor whites more employable. The surplus labor has disappeared, and today the South generally has a tighter labor market than the stagnant Manufacturing Belt, where unemployment often runs high.

Reserve labor's evaporation in the Deep South is admittedly not the whole story. High manufacturing growth rates in the South have operated from the demand side to raise wages. Also, the mix of manufacturing industry within the South has shifted away from heavy reliance on low-wage, labor-intensive industries (textiles and apparel primarily) to a better balance of high-wage and low-wage industry. All of these factors—reduced surplus labor, strong manufacturing growth, and improved industry mix—have pushed up southern wages. It would be surprising indeed if the hourly wage gains did not help cause the South's income gains.

Wage declines in the West should also have affected income. The West's relative wage decline was partly the result of the national average's being pushed up. But cost-of-living declines in the West (our next hypothesis) seem to have helped push the West's relative wages down. That idea can explain why the other high-wage region, the Great Lakes, did not also move closer to the national

average wage rate. Unlike the West, the Great Lakes had no significant freight cost reductions to push down its cost of living.

Despite these arguments for the wage hypothesis, other things point to a weakening of the wage influence between 1950 and 1987: the effect of wage rate changes could be less than the magnitude of the changes would suggest. In the preliminary evidence on wages that we summarized, the wage-income relationship was still strong in 1969 (the +.73 correlation), which is about when southern wages began to climb more rapidly. But by 1979, the relationship was much weaker (+.35 correlation), if still significant. And in 1987 no relationship between wages and income could be found for the high-wage states. The low-wage states still displayed the relationship, but it was weaker than in 1950.

What happened? Seven of the 8 states with the highest 1950 wages were in the West, and 6 of these were among the twelve with the highest income. Then the West's relative wages fell. By 1987, only 2 of the 8 states with the highest wages were western, and neither of these was among the income leaders. In other words, relative wage decline in the West pretty much destroyed the relationship between *high* wages and *high* income. Only the low-wage end of the relationship remained.

Meanwhile, the wage hypothesis is that per-capita income increases are high where increases in hourly manufacturing wages are high.

WESTERN TRANSPORTATION COST REDUCTIONS

In the West, especially the Far West, declines in relative income may have been caused by a transportation factor: shrinkage of the transport cost element in the price of manufactured goods sold in the West. The hypothesis is that this shrinkage caused cost-of-living declines, which caused wage-salary declines, which caused income declines. This chain of causation is suggested by

- the West's top-ranking per-capita income position in 1950 (table 1.1),
- the West's bottom-ranking per-capita manufacturing employment position in 1950, a position that rules out the possibility that the high income resulted from proportionately high manufacturing employment,
- the West's sharp 1950-87 decline in relative per-capita income (table 1.1), a decline that took the West from 114 percent to 105 percent of the U.S. average,
- the West's commensurate decline in hourly wages, which went from 114 percent to 103 percent of the U.S. average, and
- the Manufacturing Belt's 1950–87 wage stability (no change in Northeast, slight gain in Great Lakes), stability that suggests that the West's wage decline was not simply the result of the U.S. average wage rate's growing faster than wage rates in the Manufacturing Belt and the West.

The most remarkable relative decline in income was in the Far West (Nevada, Washington, Oregon, and California). The Far West stood at 121 percent of U.S. per-capita income in 1950; by 1987 the Far West had fallen to 111 percent of the U.S. level. For the Far West's four states, 1950 and 1987 per-capita income as a percentage of the national average were as follows:

State	1950	1987
Nevada	132%	106%
California	124	115
Oregon	110	90
Washington	114	101

In 1950 the overall West led the four census regions in per-capita income. This lead existed despite the West's being less urbanized—and far less industrialized—than the Manufacturing Belt (Northeast plus Great Lakes). Some of the West's high income was attributable to the region's ranking right behind the Manufacturing Belt in urbanization and to the West's being correspondingly weak in agriculture. Much of the remaining western income advantage was probably due to the West's being highest in median years of schooling and lowest in black population percentage among the nation's regions.

But a transportation factor was probably at work too. The West's low levels of industrialization were forcing it to import most of its manufactured goods from the distant Manufacturing Belt and other eastern locations. The 1950 per-capita manufacturing employment levels for the regions and subregions we have been observing are as follows:[39]

Region	1950 Manufacturing/Population
Deep South	0.061
Peripheral South	0.064
Northeast	0.143
Great Lakes	0.141
Plains (West North Central)	0.045
West	0.041

The West's need to import manufactured goods from the East (mainly from the Manufacturing Belt) seems to have raised the West's 1950 cost of living. No hard evidence exists to support this point, but the West's high 1950 wage levels certainly suggest a high cost of living. And at the time, many observers considered prices on manufactured goods to be higher in the West, especially in the Far West. The high western prices would logically have resulted from shipping costs embodied in those prices. The large shipping cost element in western manufactured goods resulted from a combination of (*a*) relatively high levels of imports from the East, and the resulting need to pay long-distance transportation costs on a high proportion of the goods sold, and (*b*) exceptionally high transportation costs on the imported items.

Shipping costs were exceptionally high on western imports (vis-à-vis, say, Plains imports) for two reasons. First, the goods had to be shipped longer distances. Transportation costs are proportional to distance. And the West is more distant from the Manufacturing Belt than is any other region. Second, trucks and trains taking goods to the West—the Far West in particular—had to get past as many as three mountain barriers: the Rockies, the Sierra-Cascade range, and the Pacific Coast range. Railroads and highways had to zigzag around mountains and along valleys whose directions were not the desired ones; ground distance was unusually high relative to air distance. Trains and trucks also had to climb over high-elevation passes, and the trains had to stop and take on—then later remove—extra locomotives.

In theory, then, the high transportation costs embodied in the prices of goods sold in the West raised the West's 1950 cost of living. The same phenomenon is seen today in Alaska. There, spatial isolation adds considerably to the price of manufactured goods and food shipped up from the lower 48 states. (In Alaska, the picture is complicated by long, harsh winters and a high male/female ratio, both of which require paying relatively high wages in order to attract and hold workers.) The nonwestern regions did not have this shipping cost problem: they were closer to (or else in) the East and produced more of their own goods.

The higher cost of living, still in theory, forced western employers to pay higher wages and salaries so as to enable workers to match eastern living standards. These types of compensatory wage adjustments can be seen today in such high-cost-of-living places as Anchorage and New York City. The higher wages and salaries apparently raised the West's per-capita income.

Since 1950, developments on two fronts have reduced the transport cost element in western consumer goods prices; the West's cost of living is no longer appreciably different from those of other nonsouthern regions. First, the West has consistently enjoyed the nation's highest manufacturing growth rates. Much of this growth results from regional branch plants built in the West to serve regional markets. The West's new manufacturing means that the West now makes a larger share of its own goods. No transcontinental shipping costs need be paid by consumers on the regionally produced goods. They can sell for less.

Second, many transportation developments have lowered shipping costs on those goods that must still be transported from the East. Trucking has benefited from the construction of the interstate system, the legalization of bigger trailers and double bottoms (tandem trailers, pioneered on the West Coast), and piggybacking. The railroads have gained efficiency by converting to diesel locomotives, adopting bigger rail cars, modernizing their rail yards, installing continuous rail, and getting rid of excess crewmen formerly required by union featherbedding practices.

Theory says that these developments lowered the West's relative cost of living. And the lower cost of living resulted in relatively lower wages and salaries. The West's average hourly wages in manufacturing went from 114 percent of the national average in 1950 down to 103 percent in 1987. The western decline in relative wages affected all western states: all of the West's 11 states were among the 15 states with the lowest 1950–90 percentage increases in manufacturing wages (production workers), and the West provided 5 of the 6 lowest wage

increases, including all 3 of the very lowest.[40] In comparison, relative wages rose slightly in the other high-wage region, the Manufacturing Belt, which had no transport-based cost-of-living decline. Lower wages and salaries in the West mean lower per-capita income. The hypothesis is, then, that reductions in the transport cost element in the price of manufactured goods sold in the West caused relative decreases in western per-capita income. The chain of causation runs from (1) transport costs to (2) cost of living to (3) wages and salaries to (4) per-capita income.

Direct measures of transport cost reductions are not available for testing the hypothesis. But a distance-from-Manufacturing-Belt variable has been used successfully as a proxy for transportation costs.[41] The idea behind using distance as a proxy is that shipping costs increase as distance increases. Starting with this variable, we can reasonably assume that the states with the highest 1950 transportation costs (the most distant states) had the biggest 1950-87 transport cost reductions: savings, like the original costs, should be proportional to shipping distance. The transport distance variable can therefore be used as a proxy for the missing transport cost reduction variable.

INDUSTRY MIX

The remaining hypotheses are less compelling and, in most instances, support fewer variables; hence they will get less discussion. But they are important and should not be judged by the amount of text devoted to them here. The eighth general hypothesis is that changes in industry mix induced per-capita income changes. Some industries provide above-average earnings; others provide below-average earnings. States that had relative increases in high-wage industries should show favorable effects on per-capita income.

The most important high-wage industrial sector is manufacturing. The high per-capita income in the Manufacturing Belt seems to be due partly to high manufacturing levels in the region. High urban percentages and correspondingly low farm population percentages are probably more important, however. At any rate, the West ranked ahead of the Manufacturing Belt in 1950 per-capita income yet was the least industrialized region of the United States. Still, wage earners in manufacturing generally earn more than wage earners in services, agriculture, and construction, so manufacturing growth is likely to bring increased per-capita income. A related possibility is that growth in just those manufacturing industries that pay the highest wages will stimulate income. Conceivably, growth in industries that pay low wages—textiles, apparel, leather, furniture—causes relative reductions in income.

The hypothesis that manufacturing growth stimulates income growth carries an important reservation. The western transportation cost hypothesis, just discussed, amounts to a counterhypothesis: that manufacturing growth depresses income. That is, fast manufacturing growth in the West seems to have lowered the price of manufactured goods, thereby lowering the cost of living, causing relative wage-salary reductions, and ultimately causing income to fall relative to the U.S. average. Considering the basic hypothesis and counterhypothesis (western

hypothesis) together, we would expect manufacturing growth to stimulate income in the South and Plains but to depress income in the West. If the counterhypothesis is valid for the West, a 1950–87 manufacturing change variable should be relatively weak and perhaps insignificant, at least until a transport cost change variable is controlled.

Preliminary evidence suggests that the cost-of-living effect of the lack of manufacturing in the West does in fact partly neutralize the high-wage effect of manufacturing in the East. The correlation between income and manufacturing is significant but low. Using state data for 1950, Perloff et al. found a "small but significant" correlation of +.33 between the median income of urban and rural nonfarm persons (combined) and the percentage of the labor force employed in manufacturing.[42] Unpublished data from Wheat's county study show a weaker but still very significant correlation (based on 2,706 cases) of +.18 between 1969 median family income and manufacturing employment as a percentage of total employment. Schwirian found a +.20 correlation between 1979 county median family income and high-wage manufacturing employment (all but textile and apparel) as a percentage of total employment.[43]

The service sector offers two competing hypotheses that call for opposite effects. The broadly defined service sector—services per se plus wholesale and retail trade and finance, insurance, and real estate—pays generally low wages. Growth in overall services should therefore lead to a relative decline in per-capita income. But growth in particular service industries that pay relatively high wages or have proportionately high professional employment might have the opposite effect. Coughlin and Mandelbaum speculate that comparatively fast per-capita income growth in New York, New Jersey, and Rhode Island during 1978–87 might have been helped by service growth, particularly in health, business, and financial services.[44]

Mining, which includes petroleum extraction, is another high-wage industrial sector. Mining expansion, often associated with exploitation of newly discovered petroleum or ore deposits, has often led to increased income. Mining contraction, often associated with resource depletion (as with oil in Louisiana and Texas and iron ore in Minnesota), has depressed income. Logically, then, relatively fast growth in mining employment should be associated with fast growth in per-capita income.

Construction, transportation, and public administration seem unlikely to have much effect on income. They are relatively small sectors and are too close to average in their hourly wages to promise important income effects. Wheat's county study did find the construction employment percentage significant: it was the weakest variable in a 20-variable equation predicting 1969 county median income.[45] Construction employment's simple correlation of -.24 described low income among construction workers. When we are dealing with *change* in construction employment rather than *levels*, however, the results could be different. The stimulating effects of economic booms, with their increase in construction employment, are likely to cancel the depressing effects of low income in the industry.

The subordinate hypotheses for industry mix are that per-capita income changes will be high where

(1) overall manufacturing employment increases are relatively high,

(2) high-wage manufacturing employment increases are relatively high,

(3) combined high-wage and mid-wage manufacturing employment increases are relatively high,

(4) low-wage manufacturing employment increases are relatively low,

(5) coal, metal, and mineral mining employment increases are relatively high,

(6) petroleum extraction employment increases are relatively high,

(7) overall mining employment increases are relatively high,

(8) high-income service industry employment increases are relatively high, and

(9) total service employment increases are relatively low.

ENERGY PRICE CHANGES

The ninth general hypothesis is that energy price changes that began in 1973 affected income in petroleum-producing states. Energy prices climbed rapidly in the 1970s in response to the 1973–74 OPEC oil embargo and the 1979 Iranian revolution. The unprecedented energy price increases of the 1970s were followed by price declines after 1981, the year energy prices peaked. The energy hypothesis is weak, because the post-1981 price declines more or less cancelled the post-1973 increases. Still, the hypothesis is worth checking out.

To some extent, the energy price change effects may be captured by some of the variables that test industry mix hypotheses. Part of any energy stimulus to per-capita income would be employment increases in oil and national gas extraction and in petroleum refining. These employment increases resulted from petroleum shortages and from associated oil and gas price increases. Similarly, energy price declines in the 1980s should be reflected in energy employment declines. Petroleum extraction and refining employment variables might therefore account for any energy effects. Indeed, so might E/P change and related variables.

But the energy price changes did not affect income only through the medium of employment changes. The price changes per se should have brought increased income to oil-producing and gas-producing states. Even without increases in energy resource production, higher prices applied to existing levels of production could generate higher per-capita income. We can therefore hypothesize that states with high mid-1970s employment in oil-gas extraction and petroleum refining will show high increases in per-capita income. (Note that this particular variable, unlike almost all the rest, measures a particular year's *level* of the indicator rather than a 1950–87 *change*.) We can also hypothesize that, after the mid-1970s variable is controlled, states with high mid-1980s energy industry employment will show low increases in per-capita income. (Before control of the mid-1970s

variable, the mid-1980s variable would tend to be positive as a proxy for the mid-1970s variable.)

The specific variables, which will be listed shortly, will use 1974 and 1982 energy employment per capita. Both years are one year after the price upturn (1973) or downturn (1981). The one-year lag gives energy employment responses to the supply and price developments a year to work themselves out. This one year for employment adjustment serves two purposes. First, the adjustment makes it harder for the energy variables to act as proxies for the industry employment change variables (industry mix variables). In doing so, the adjustment year increases the possibility that the equation will accept both industry mix variables (measuring employment effects) and energy price variables, in case both employment change effects and price change effects are operating. Second, the one-year adjustment period allows states that changed their energy employment in response to the 1973 and 1981 developments to show corresponding price effects. In other words, if Utah expanded its oil production between 1973 and 1974, the 1974 level is more likely to be reflected in Utah's 1987 per-capita income.

SHORT-RUN DEVELOPMENTS OF THE 1980S

The tenth and last general hypothesis focuses on developments in the 1980s. These developments caused income changes that differed from those of earlier periods. From 1950 until 1978, the general pattern of change in state per-capita income was one of convergence on the national average; the dominant subtrend was that of the South's catching up with the rest of the country. But, as we saw in chapter 2, the long-run trends reversed during 1978–88; state per-capita incomes diverged. To some extent, this divergence seems to have been the indirect result of the playing out of the forces that had been causing convergence. By 1970, the South's sharecropper system was essentially dead, and the Great Black Migration was over. Whereas declines in farm population (particularly sharecropper population) had been helping the South and other agricultural regions to catch up, the nation's farm population percentage had fallen from 15 percent in 1950 to 4 percent in 1976 and 3 percent in 1982. And, as seen earlier, the South had almost caught up with the other regions in education by 1980. With the formerly powerful convergence-producing forces thus weakened, new influences were able to cause divergence. These divergence-producing influences would probably have been swamped by the forces of convergence in former decades.

The two most conspicuous subtrends of the divergence trend of the 1980s were (1) the economic boom in New England, and, to a lesser extent, in the Middle Atlantic states and (2) resource-related subregional recessions in some southern and western states. Falling energy prices, petroleum depletion, and copper ore depletion had adverse effects in the South and West. Regarding the Northeast's boom, Coughlin and Mandelbaum list ten states that diverged appreciably from national average per-capita income during 1978–87 through relative *increases* in per-capita income.[46] Nine of the ten states are in the Northeast. (The tenth state was Florida.) Two other Northeast states—Vermont and Maine—also gained in

relative per-capita income but converged on the national average, because they started below the average.

The Northeast's boom and New England's leadership in that boom are reflected in regional and subregional unemployment rates. The civilian unemployment rate for the United States in 1987 was 6.2 percent. But there were subordinate unemployment rates of 3.3 percent in New England, 4.9 percent in the Middle Atlantic states, 6.7 percent in the Midwest (combined Great Lakes and Plains), 6.8 percent in the South, and 6.3 percent in the West. The point is that an economic influence—the regional boom—that caused divergence registered in the 1987 regional unemployment rates.

The other main subtrend causing divergence was the resource-related recessions in certain southern and western states. Coughlin and Mandelbaum list another ten states that appreciably diverged from national average per-capita income during 1978–87 through relative *decreases* in per-capita income. Four of these ten states—Texas, Louisiana, Oklahoma, and New Mexico—were among the six leading oil-producing and gas-producing states of the lower 48 in 1982. All four suffered from the post-1981 oil price declines, and all but New Mexico additionally suffered from 1982-87 production decreases.[47] (Production was virtually unchanged in New Mexico.) The fifth-ranked oil-producing state, Wyoming, also declined in per-capita income but wound up closer to the national average: it went from 117 percent to 89 percent of national average per-capita income. The last of the six oil production leaders, California, had a more diversified economy and did not change in relative per-capita income between 1982 and 1987.

Two other states among the ten most downwardly divergent likewise experienced resource-related recessions. The states were Montana and Utah; the resource was copper. World copper price declines and depletion of higher-grade ore caused the problem. In Montana, tonnage went from 72 million in 1982 to 17 million in 1985, the only post-1984 year for which figures are available. In Utah, production declined from 208 million tons in 1982 to 187 million tons in 1984.[48] (Recoveries in copper production occurred after 1987.)

Just as the upward divergence caused by the Northeast boom was reflected in 1987 unemployment rates, so was the downward divergence caused by the oil-gas and copper recessions. Seven resource-economy states had appreciable percentage point declines in per-capita income as a percentage of the national average during 1978–87. Six of these diverged, and one (Wyoming) converged. The seven states, ranked high to low by percentage point decline in relative per-capita income, had these 1987 unemployment rates: Wyoming, 8.6 percent; Montana, 7.4 percent; Louisiana, 12.0 percent; Utah, 6.4 percent; Oklahoma, 7.4 percent; New Mexico, 8.9 percent; and Texas, 8.4 percent. For all seven states, the 1987 state unemployment rate exceeded the national unemployment rate of 6.2 percent.

This analysis suggests that the 1987 unemployment rate can be used to measure important short-run influences that might not be picked up by the long-run change variables. That is, a state's 1987 unemployment rate can serve as a proxy for recent economic developments that might have caused particular states to have exceptionally high or low unemployment in 1987. Secondarily, the unemployment rate can describe the effect of unemployment per se on 1987 per-capita income,

hence on 1950–87 change in income. This secondary role, however, is likely to be minor. The main unemployment rate variable, which measures the 1950-87 *change* in the unemployment rate, is better suited for registering the effect of unemployment rates per se. The 1987 *level* of the unemployment rate is theoretically best for proxy purposes, because in the proxy role it is able to describe *changes* in subregional economic conditions that affect per-capita income.

The hypothesis, then, is that per-capita income gains were high in states that had low unemployment rates in 1987. Two subordinate hypotheses are that per-capita income gains were

(1) high in states where subregional economic booms occurred in the last decade or so of the 1950–87 period and

(2) low in states where resource-related recessions occurred in the last decade or so of the period.

REGIONAL INTERACTION HYPOTHESES

Some additional hypotheses overlap the preceding ones, hence cannot be counted as truly separate hypotheses. These additional hypotheses concern the possibility that some of the ten basic hypotheses are valid only in, or else outside, the South. Or, the hypotheses may have greater validity (i.e., the variables may be more significant) in one region than in the other, where the regions are South and non-South; separate southern and nonsouthern calibrations could be desirable for certain variables. For example, the South had bigger rural-urban wage differences, so rural-to-urban population shifts may have been a stronger income stimulus in the South. A related possibility is that different variables within the same group, such as the rural-urban group, may better describe the basic relationship in different regions. As a case in point, and to anticipate, change in urban population percentage best describes the effect of rural-urban mix changes in the South, but farm variables prove to be the best measures of rural-urban mix outside the South.

The above discussion alludes to the use of regional dummy variables (slavery dummy, nonslavery dummy) as multipliers in interaction variables, such as SLAVExPOP%METRO. The dummy multiplier "dummies in" those states where the variable's effect is expected to be strongest; the other states (for which the dummy is valued at zero) acquire values of zero. The central question in this application of regional dummies is: what regions should be delineated for purposes of identifying differences in the regional effects of certain factors?

The 48 states were divided into slave (South) and nonslave (non-South) for cogent reasons. Per-capita income change for 1950–87 has been dominated by the catching-up trend, wherein the states of American's most distinctive region—the South—have been closing in on the nonsouthern states. The basic cause of this catch-up trend has been the collapse of a distinctively southern institution: the sharecropper-tenant farmer system. Sharecropping's collapse ramifies into

various subordinate causal influences. These include the subordinate catching-up trends in (1) urban population percentage, (2) educational attainment, (3) population percentage of workers, (4) population percentage of whites, and (5) hourly wage rates. In all five of these trend areas, the South was the distinctive region in 1950; it was the region that differed appreciably from the norm observed by other regions. And all five trends were related to the collapse of sharecropping. Or, one could say, all five trends embodied the legacy of slavery: slavery was ultimately responsible for the South's unfavorable standing (relative to per-capita income) in the five areas. Since slavery was a southern institution, this common theme once more points to the South as a distinctive region for analytical purposes.

This is not to deny that per-capita income changes were uneven among the nonsouthern states. The nonsouthern changes *were* uneven. In general, the western states—highest in 1950—had the biggest declines in relative income. The Plains states—lowest in 1950—tended to fare the best. And within regions and subregions, some states changed more than others.

Why were the nonsouthern changes uneven? One reason, perhaps the most important, is that the farm population decline was uneven outside the South. Farm population declined the most in the most agricultural states and in the most agricultural subregion, the Plains. In the arid mountain states of the West, where agriculture is more oriented towards sheep and cattle ranching, new farm technology had less effect, and there were proportionately fewer people in agriculture to begin with; the farm decline had less influence on western income. A second reason for uneven change in nonsouthern per-capita income was uneven growth in the black population percentage. The blacks who migrated from the South went mainly to metropolitan areas (hence metropolitan states) and to the states located closest to the South. New York got more black migrants than Vermont; Illinois got more than Minnesota and Wyoming. A third reason for uneven change in nonsouthern income may have been the uneven effect of Manufacturing-Belt-to-hinterland-state transportation cost reductions. These reductions were greater in the West, and particularly in the Far West. A fourth reason for uneven nonsouthern income changes is probably that the population percentage of divorced women (low income) climbed most rapidly in the more metropolitan states.

So, nonsouthern influences—not just southern influences—were definitely at work. But the nonsouthern developments and influences were secondary from a national perspective. The essential point here is that the most important causal factors had strong southern roots. This fact creates the possibility that the factors operated differently—and perhaps had more effect—in the South than elsewhere. Another possibility is that some factors, though rooted in southern developments, had more effect outside the South.

Change in black population percentage is one such factor—a factor that may have had more effect outside the South. The *relative* change in black population was generally much greater in the nonsouthern states than in the South. The black population percentage declined 16 percent in the Deep South during 1950-87, but it increased 116 percent in the Northeast, 102 percent in the Great Lakes, 63 percent in the Plains, and 97 percent in the West. Moreover, because of the

strong correlation between educational advances and black population changes within the southern block of states, there is a good possibility that education variables duplicate the effect of black changes in the South but not elsewhere. A NONSLAVExPOP% BLACK variable could be more significant than the basic variable, POP%BLACK.

These considerations justify the choice of South and non-South as the regional delineation to be used in searching for regional differences in the effects of certain causal influences. Except in the case of retiree migration, the causal influences in question are ones that relate ultimately to slavery and immediately to slavery's effects on black education and black employment.

A list of subordinate regional interaction hypotheses at this point would be excessively long and would cover ground previously covered. A general hypothesis will suffice. The general hypothesis is that factors related to slavery's legacy will sometimes have more or less effect in the South than in the non-South. Although "South" and "non-South" are used for convenience, the census definition of the South (16 states) will not be used. Instead, a more relevant definition will be used. For purposes of constructing South and non-South dummy variables, the South is defined as the 15 slave states—the census South minus West Virginia and Oklahoma and plus Missouri.

NOTES

1. Leonard F. Wheat, *Labor Force Participation and Unemployment in American Counties* (Washington, D.C.: Economic Development Administration, 1982), Appendix A, "Determinants of Median Family Income," A-6.

2. Kent P. Schwirian, *Determinants of County and City Median Family Income and Poverty Rates* (Washington, D.C., Economic Development Administration, 1991), 15.

3. U.S. Department of Agriculture, Economic Research Service, *Rural People in the American Economy*, Agricultural Economic Report 101 (Washington, D.C., October 1966), 16.

4. Wheat, *Participation*, table 3.4.

5. U.S. Department of Commerce, Bureau of the Census, *Residents of Farm and Rural Areas*, Current Population Reports P-20, no. 446 (Washington, D.C., 1991), table A.

6. U.S. Department of Commerce, Bureau of the Census, *Statistical Abstract, 1989* (Washington, D.C., 1989), table 32.

7. The 1950 percentage is from the 1956 *County and City Data Book*, table 1. The 1987 percentage is from Bureau of Census, *Residents of Farm and Rural Areas* (note 5). The 1987 urban percentage is less than the 1987 metropolitan percentage because the metropolitan percentage includes persons living in rural parts of metropolitan areas. The urban percentage includes all people living in urban places, whether in metropolitan or nonmetropolitan counties.

8. See note 7 for sources. The regional configuration used in the tables in this chapter differs slightly from that used for per-capita income in table 1.1. Specifically, Delaware and Maryland have been moved from the BEA Mideast region (used in table 1.1) to the Census Bureau's South, where they go in the Peripheral South subregion. The reason for the change is that the tabulations in chapter 4 use Census Bureau sources that aggregated

states into the census configuration of four regions and nine divisions. To permit continued comparison of Deep South and Peripheral South characteristics, these two unconventional regions use hand-calculated statistics. In the interest of simplicity, the Northeast and West are not broken into subregions in this chapter; they are sufficiently homogeneous that nothing vital is lost. The Northeast uses the census definition. The Northeast and Great Lakes regions plus Delaware and Maryland constitute the Manufacturing Belt.

9. Wheat, *Participation*, A-6.

10. Schwirian, *Income and Poverty*, 21.

11. Nicholas Lemann, *The Promised Land: The Great Black Migration and How It Changed America* (New York: Knopf, 1991), 47.

12. Ibid.

13. Gavin Wright, *Old South, New South: Revolutions in the Southern Economy Since the Civil War* (New York: Basic Books, 1986), 80.

14. 1956 *County and City Data Book*, table 1, and *1980 Census of Population, General Social and Economic Characteristics, U.S. Summary*, report no. PC80-1-C1, tables 83 (regions) and 239 (divisions and states).

15. Wheat, *Participation*, A-6.

16. Schwirian, *Income and Poverty*, 13.

17. 1956 *County and City Data Book*, table 1 (separate population and employment totals), and 1989 *Statistical Abstract*, table 623 (E/P precalculated).

18. Wright, *South*, 96.

19. Ibid., 177-86.

20. Wheat, *Participation*, A-9.

21. Schwirian, *Income and Poverty*, 23.

22. Wheat, *Participation*, 11-3.

23. Eli Ginzberg, "Youth Unemployment," *Scientific American* 242 (May 1980), 47.

24. Wheat, *Participation*, 113.

25. Nicholas Lemann, "The Origins of the Underclass," *Atlantic Monthly* 259 (June 1986), 35. Cf. Albert Rees, "An Essay on Youth Joblessness," *Journal of Economic Literature* 24 (June 1986), 623.

26. Ginzberg, "Youth," 46.

27. Ibid., 47.

28. *1950 Census of Population, Characteristics of Population, United States Summary*, table 59, and 1989 *Statistical Abstract*, table 26.

29. President's National Advisory Commission on Rural Poverty, *The People Left Behind* (Washington, D.C., 1967), 6.

30. Wheat, *Participation*, table A-1. The partial correlation is from an unpublished tabulation.

31. Schwirian, *Income and Poverty*, 24, 57.

32. William H. Crown, Phyllis H. Mutschler, James H. Schulz, and Rebecca Loew, *The Economic Status of Divorced Older Women* (Waltham, Mass: Brandeis University, Policy Center on Aging, 1993).

33. William Goode, *The Family* (Englewood Cliffs, N.J.: Prentice-Hall, 1982); and Roy Marshall, *The State of Families* (Milwaukee: Family Service America, 1991).

34. William H. Crown, "State Economic Implications of Elderly Interstate Migration," *The Gerontologist* 28 (1988), 533-39; B. Richard, "Old Money: An Influx of Retirees Pumps New Vitality into Distressed Towns," *Wall Street Journal*, August 5, 1988; and W.

Hass III and W. Serow, "The Influence of Retirement In-Migration in Local Economic Development," Final Report to the Appalachian Regional Commission (Asheville, N.C.: University of North Carolina, 1990).

35. The age statistics in this paragraph are from the 1956 *County and City Data Book*, table 1, and the 1983 *County and City Data Book*, table A.

36. Wright, *South*, 30.

37. Lynn E. Browne, "Narrowing Regional Income Differentials," *New England Economic Review* (September-October 1980), 44.

38. 1956 *Statistical Abstract*, table 265, and 1989 *Statistical Abstract* table 1267. For this variable, the regional and national averages are simple rather than weighted averages of state figures.

39. Computed from state data in Leonard F. Wheat, *Regional Growth and Industrial Location* (Lexington, Mass.: D. C. Heath, 1973), table 1.2.

40. Oliver J. Blanchard and Lawrence F. Katz, "Regional Evolutions," *Brookings Papers on Economic Activity* 1 (1992), figure 4.

41. Leonard F. Wheat, "The Determinants of 1963-77 Regional Manufacturing Growth: Why the South and West Grow," *Journal of Regional Science* 26 (1986), 638, 655.

42. Harvey S. Perloff, Edgar S. Dunn, Jr., Eric E. Lampard, and Richard F. Muth, *Regions, Resources, and Economic Growth* (Baltimore: Johns Hopkins Press, 1960), 528.

43. Schwirian, *Income and Poverty*, 18.

44. Cletus C. Coughlin and Thomas B. Mandelbaum, "Why Have State Per Capita Incomes Converged Recently?" *Federal Reserve Bank of St. Louis Review* 70, no. 5 (September-October 1988), 30.

45. Wheat, *Participation*, table A-1.

46. Coughlin and Mandelbaum, "Incomes," 26.

47. *Statistical Abstract*, table 1189.

48. Ibid., table 1203.

5

METHODOLOGY, VARIABLES, AND CORRELATION FINDINGS

This study uses correlation-regression analysis. We have two basic models of 1950–87 per-capita income change in the 48 contiguous states. The models test the ten general hypotheses from chapter 4 (e.g., the education hypothesis) and a much larger number of specific ones (e.g., the *college* education hypothesis). The models also develop evidence that allows us to rank the determinants of income change according to importance. Each basic model has many variations. The basic models are defined by their dependent variables; the variations involve differences in the number and identity of the independent variables. In this chapter we describe the basic models, identify the variables tested, and discuss the correlation findings. Some of the finer details of our methodology bear only on the regression findings, not on the correlation findings. We reserve those details for chapter 6, where we discuss the regression findings.

METHODOLOGY AND VARIABLES

Specific considerations determined (1) our choice of 1950–87 as the study period, (2) the geographic coverage of the models, (3) the choice of dependent variables, (4) the design of the independent variables, and (5) the decision to discuss correlation as well as regression findings.

The Study Period

The 1950 starting point of the 1950–87 study period embodies census considerations: 1950 was the year of the first postwar census. Two things influenced our decision to start with this particular census. First, a 1950 starting point avoids distortions caused by possibly profound but cross-sectionally hard-to-

explain income changes caused by World War II. Second, 1950 was the first census after the large-scale adoption of the mechanical cotton picker began. The mechanical cotton picker, later complemented by chemical weeding (herbicides) and other advances in farm technology, was largely responsible for the collapse of the sharecropper-tenant farmer system—the institution that had depressed southern income since the Civil War. In other words, 1950 was the approximate starting date of the technological revolution in southern (and northern) agriculture.

The study period's 1987 ending point was also chosen primarily for data reasons: 1987 was the latest year for which BEA per-capita income figures were available at the time the study was begun. But 1987 also offered a second advantage. The 1987 unemployment rate (6.2 percent) is closer to the 1950 unemployment rate (5.3 percent) than is that of any other available recent year. Over the 1980–86 period, annual average unemployment ranged from 7.0 percent (1986) to 9.7 percent (1982). The choice of 1987 therefore minimized the amount of possible business cycle distortion in the findings.

Geographic Coverage

The geographic scope of the study is limited to the 48 contiguous states; the sample excludes Alaska, Hawaii, and the District of Columbia. Alaska and Hawaii were omitted for two compelling reasons. First, they were not states in 1950; hence they lacked data for some of the variables. Second, a possibility—indeed, a likelihood—existed that their extreme values for certain variables would create outlier problems and consequently biased findings.

The outlier problems are especially serious for Alaska. Alaska's 1950 per-capita income was far above that of any other state or territory. The explanation lies in the premium wages needed to attract workers to Alaska. Premium wages were needed because of (1) the harsh winter climate, (2) spatial isolation and resulting high personal expenses associated with migration and travel, (3) a high cost of living, which resulted from both high wages and high transportation costs on imported food and manufactured goods, and (4) a relative lack of unmarried women, which made Alaska somewhat unattractive to men contemplating moving there. Between 1950 and 1987 Alaska declined from 154 percent of national per-capita income (far ahead of second-ranked Delaware's 136 percent) to a fifth-ranking 119 percent of national per-capita income. This Alaskan change of 35 percentage points dwarfs the next biggest change, Mississippi's 15-point increase (from 52 to 67 percent), and makes Alaska an extreme outlier—a case that would have disproportionate influence on cross-sectional correlation-regression results. Yet the causes of Alaska's income change differ greatly from the causes in other states. Including Alaska would have seriously distorted the study findings.

Including Hawaii would not have been as dangerous but would still have caused problems. Hawaii would be an outlier, and one with anomalous behavior at that, for at least one important variable: the transport cost reduction variable. More important, Hawaii's 1950–87 per-capita income change was probably influenced heavily by something that had little effect in other states—and that

would be hard to capture with an available variable. During 1950–87, Hawaii's per-capita income rose from 92 percent of the national level to 101 percent. Much, perhaps most, of this gain resulted from a boom in tourism, a boom facilitated by the tremendous expansion of air travel after 1950. This air-and-tourism influence is probably unique among the states. What other state that had a substantial increase in relative income can attribute most of that increase to tourist industry growth based on lower air fares? In short, putting Hawaii in the study would have let a singular influence dilute the influences that operate in the 48 contiguous states. A model's ability to register the main-line influences would have been seriously impaired.

As for the District of Columbia, it is not a state. It is not even a metropolitan area. It is a metropolitan central city. As such, it is another outlier for many variables. For example, its per-capita income was 145 percent of the national average in 1950. That put it well above the nearest state (Delaware, 136 percent) and only slightly below Alaska (154 percent). In 1987 the District ranked behind only Connecticut and New Jersey. The District is also an outlier for the farm–urban variables and the black population percentage. Its status as a central city made it 100 percent urban both years, so it had the least change—no change at all—in farm population percentage, urban population percentage, and metropolitan population percentage. Central city status also gave it the biggest black population percentage increase over the period. Researchers who include the District in state cross-sectional analyses err. They fail to recognize that the consideration that motivates the Census Bureau to include it in state tables (100 percent inclusiveness of the national figure for any statistic) is not relevant where cross-sectional analysis is concerned.

Dependent Variables

The two basic models are (1) the Percentage Change model and (2) the State/Nation model. Each uses state per-capita income change as the dependent variable, but they measure change differently. The Percentage Change model uses the 1950–87 percentage increase in per-capita income; the state value is *not* expressed as a percentage of the national value before change is computed. For this model, both 1950 income and 1987 income are measured in 1987 dollars. The dependent variable is

(1) INCOME/POP:% = 1950–87 percentage increase in state per-capita income, or [(1987 - 1950)/1950] x 100.

The State/Nation model measures state per-capita income as a percentage of national per-capita income before change is computed. That is, each *state's* per-capita income is divided by the *nation's* to get a State/Nation ratio, which is then converted to a percentage (i.e., multiplied by 100). The 1950 percentage is then subtracted from the 1987 percentage to get the absolute change (not necessarily an increase) in per-capita income. The dependent variable is

(2) INCOME/POP:S/N = 1950–87 absolute change in the ratio of state to national per-capita income (S/N, expressed as a percentage), or 1987 S/N - 1950 S/N.

Our reason for using two models—two dependent variables—is that there was no way to determine in advance what method of measuring change would prove most sensitive, or best able to capture the influences that determined change. We had four possible ways of measuring change: (*a*) two ways of measuring 1950 and 1987 income levels interacting with (*b*) two ways—relative and absolute—of measuring change. INCOME/POP:% seemed to us to be the most straightforward or basic measure, so we used it for our first model. We chose INCOME/POP:S/N for the second model because this variable has the least resemblance to the first: it differs in both the way it measures income and the way it measures change.

Independent Variables

Two aspects of the independent variables require explanation: (1) structural aspects and (2) substantive aspects.

Structural Aspects

Each model works with 53 independent variables. (These are all tested, but most never get into any of the final models.) All but 9 of the 53 variables measure change. The exceptions are (1) the 1975–80 net migration rate for persons 60 and over, (2) 1974 energy industry employment per capita, (3) 1984 energy industry employment per capita, (4) a dummy identifying the 11 states most affected by energy price declines occurring after 1981, (5) a dummy identifying the 12 states most affected by farm price declines occurring after 1981, (6) a dummy identifying 19 states having natural amenities that attract retirees, (7) the 1987 unemployment rate, (8) a regional dummy identifying the 15 slave (southern) states, and (9) a regional dummy identifying the 33 nonslave (nonsouthern) states. The two regional dummies are both used as multipliers in some interaction variables, so the second dummy is not redundant (despite the -1.00 intercorrelation between the two dummies). The nine exceptional variables just named are identical in the Percentage Change and State/Nation models. A tenth variable, the transport cost reduction variable, is also identical in the two models; but it is not listed above, because it *does* measure change (indirectly).

The remaining 43 variables are substantively identical but structurally different in the two models, subject to a minor qualification. Specifically, each of the 43 except POP%BLACK has two versions: one for the Percentage Change model and one for the State/Nation model. Substantively, the two versions are identical: they measure the same thing. For example, the two variables labeled POP%METRO both measure 1950–87 change in the percentage of metropolitan area residents in the overall population. But the two variables differ in their structures—in the way change is computed. Each version computes change in the

same way that its model's dependent variable computes change. The Percentage Change version uses the 1950–87 percentage change in the independent variable's state value. The State/Nation version separately calculates the 1950 and 1987 ratios of state-to-U.S. values (essentially percentages); then it calculates the 1950–87 absolute change (1987 minus 1950) in each variable's state/nation ratio.

POP%BLACK, the 1950–88 change in the black population percentage, used State/Nation values in both models. (Because 1987 data were not available, the variable uses 1988 data.) The Percentage Change version, originally tested, had an unbelievably low simple r of -.02. The problem was that POP%BLACK's percentage increases were badly infected with the small-base effect in many nonsouthern states. Huge percentage increases were being measured in states that in 1950 had only tiny black percentages, less than 1 percent—for example, 0.04 percent in North Dakota, 0.11 percent in South Dakota, 0.12 percent in Vermont, 0.14 percent in New Hampshire, 0.18 percent in Idaho, 0.21 percent in Montana, 0.40 percent in Utah, and 0.47 percent in Minnesota (all less than one-half of 1 percent). Black in-migration to these states after 1950 caused extremely high percentage increases, because the percentages were being computed from small—often minuscule—bases. The affected states became statistical outliers that destroyed the statistical relationship POP%BLACK was supposed to assess. This problem does not affect the State/Nation version of POP%BLACK, because that version computes the 1950-88 increases by subtracting, not by dividing. Substituting POP%BLACK's State/Nation version for the Percentage Change version in the Percentage Change model raised r from -.02 to -.61. An interaction variable, SLxPOP%BLACK, multiplies POP%BLACK by a slave state dummy (SLAVE). Because the former slave states (southern states) do not have the small base problem, we used the Percentage Change version of POP%BLACK for SLxPOP%BLACK in the Percentage Change model.

Substantive Aspects

Table 5.1 lists the study's independent variables and gives the sources of the data used to compute them. The table groups the variables under seven headings: (1) rural-urban mix, (2) educational attainment, (3) worker-nonworker mix, (4) high-income and low-income groups, (5) industry mix, (6) miscellaneous (transport cost reductions, wage rates, energy status, farm crisis status, and amenities), and (7) regional dummies and regional interaction variables. Each of the 43 *change* variables among the 53 independent variables listed measures change between two years. When possible, the years are 1950 and 1987. But other periods are used if 1950 or 1987 data are unavailable; we view the variables that use other periods (other than 1950–87) as proxies for their unavailable 1950–87 counterparts.

To anticipate, our longest and best model has an R of .943 and an adjusted R^2 of .859; the model explains 86 percent of the variance in per-capita income change. Recall that most of our leading variables measure 1950–80 change. They thus cover only 30 of the 37 years in (or 81 percent of) the 1950–87 study period. This limitation reduces the model's explanatory power.

The data for the first and last years of a change variable's coverage period

generally came from two different sources. And when the data from those sources were in absolute form (e.g., total farm population), we also used a third source to convert the one-year values to relative form (per-capita or percentage of population). That third source was a BEA computer printout showing state time-series population figures. The 1950 hourly wage rates were converted to 1987 dollars before change was measured. This conversion entailed using still another source: the 1989 *Economic Report of the President*, in which table B-58 has the consumer price index (CPI) time series. (This same CPI time series was used for preparing INCOME/POP:%.)

The Correlation-Regression Approach

This study combines correlation analysis with regression analysis; we use both. We would be among the first to agree that simple *r*'s can be misleading, since they hold nothing constant. (Partial *r*'s have one or more variables controlled.) A high, or at least significant, *r* can easily result from a proxy relationship or from a combination of proxy relationships. A proxy relationship may exist when the significant variable has a high intercorrelation with another significant variable. (The prefix *inter* in *intercorrelation* signifies that *r* measures the relationship between two independent variables or between two dependent variables, not between an independent and a dependent variable.)

An example of a proxy-based *r* is the +.56 *r* achieved by a right-to-work dummy in a manufacturing growth rate study. This *r* rested on not just one but four proxy relationships—four favorable intercorrelations with genuinely significant variables. The +.56 moved steadily in the negative direction as the genuine variables were controlled one by one: the +.56 simple *r* became a -.21 (wrong sign) partial *r*.[1]

The intercorrelation problem, incidentally, goes beyond spurious correlation (false significance). Insignificant simple *r*'s can result from *unfavorable* intercorrelations when variables are genuinely (latently) significant. Insignificant *r*'s can also result when initially insignificant variables are relatively unimportant and are overshadowed by more powerful ones: one or more variables must sometimes be controlled before the significant relationship shows up. Examples appear in this study.

But being potentially misleading does not make *r*'s worthless. Sometimes they are the only evidence available; if nothing else, they are food for hypotheses—a reason for further investigation. More important, *r*'s can provide valuable *supplemental* evidence, to be considered along with other evidence, for judging which of several causal influences is most important. When several variables are being compared, the one with the highest *r* probably represents the strongest influence. (This statement assumes that the researcher has done his homework and has provided suitable variables for all of the most important influences. Otherwise the strongest variable may rank highest because of a proxy relationship involving a missing variable.) Similarly, when two highly intercorrelated variables both

have high *r*'s and the researcher is uncertain about who is a proxy for whom, the variable with the higher *r* is likely to be the one that speaks for itself.

And *r*'s have still other uses. Intercorrelations help identify proxy relationships and partial proxy relationships. (Here "partial" means that a variable speaks partly for itself and partly as a proxy for one or more other variables. In this study the education variables are partial proxies for the black population percentage and, initially, for urbanization as well.) Intercorrelations involving regional dummy variables (SLAVE in this study) can be used to learn whether other independent variables tend to have high or low values in certain regions. We pay special attention to this application.

Beyond simple *r*'s, partial *r*'s—these hold one or more other variables constant—can be extremely helpful. They can (*a*) demonstrate that a certain simple *r* leans heavily on a proxy relationship and (*b*) bring out latent significance in a variable when that significance is hidden or minimized by an unfavorable proxy relationship.

The preceding comments do not imply that *r*'s outweigh regression findings in importance. Our point, rather, is that the two types of findings are complementary. We have far more regression findings than correlation findings, because we have many equations per basic model. The regression findings weigh more heavily in our analysis, not only because there are more of them but because they hold things constant.

CORRELATION FINDINGS

Both the Percentage Change model and the State/Nation produced useful correlation findings. We will discuss the Percentage Change findings in far more detail, because the Percentage Change model proved to be the more sensitive in detecting relationships. The State/Nation findings tend to duplicate the first set of findings but carry generally lower *r*'s.

Percentage Change Model

The Percentage Change model's dependent variable (INCOME/POP:%) is, once again, the 1950–87 percentage increase in state per-capita income. Table 5.2 is a correlation matrix for selected variables from this model. The variables shown are those with the best simple correlations (nothing held constant) plus a few others that have latent significance. The latent-significance variables appear in some of chapter 6's regression equations. We will look separately at (*a*) correlations (*r*'s) between the dependent variable and the independent variables, (*b*) intercorrelations between SLAVE and other independent variables, and (*c*) other intercorrelations of special interest. By one-tail test, the 5 percent and 1 percent significance levels for an *r* based on 48 cases are .24 and .33.

Table 5.1
Independent Variables and Their Sources
(Expected Signs in Parentheses)

	RURAL-URBAN MIX
POP%METRO	1950–87 change in percentage of state population living in metropolitan areas. Sources: 1950—*County and City Data Book, 1956*, table 3, p. 346. 1987—*Statistical Abstract, 1989*, table 35. (+)
POP%URBAN	1950–80 change in percentage of state population living in urban places. Sources: 1950—*County and City Data Book, 1956*, table 1. 1980—*1980 Census of Population*, report PC80-1-A1 (U.S. Summary), table 13. (+)
POP%FARM	1950–80 change in percentage of state population living on farms. Sources: 1950—*1950 Census of Population*, v. 2, table 24. 1980—*State and Metropolitan Data Book, 1986*, table C. (-)
POP%TENANTFARM	1950–80 change in percentage of population living on tenant farms, including sharecropper farms. Sources: Same as for POP%FARM. (-)
POP%RURNONFARM	1950–80 percentage change of state population living in rural areas but not on farms. Sources: Same as for POP%URBAN and POP%FARM (rural-nonfarm percentage is 100 - urban percentage - farm percentage). (-)
ACRES/FARM	1950–80 change in average acres per farm. Sources: 1950—*Statistical Abstract, 1954*, table 746. 1980—*Statistical Abstract, 1989*, table 1079. (+)
FARM%xACRES	POP%FARM x (ACRES/FARM) (interaction variable). So that the weaker variable, ACRES/FARM, will have less weight in the interaction formula, ACRES/FARM is recomputed as a 1980/1950 *ratio* instead of a percentage increase for purposes of this variable (e.g., a 30 percent increase becomes 1.3). (-)
TENANT%xACRES	POP%TENANTFARM x (ACRES/FARM) (Interaction variable). ACRES/FARM is again computed as a 1980/1950 ratio. (-)
	EDUCATIONAL ATTAINMENT
MD-YRS-SCHOOL	1950–80 change in median years of school completed, persons 25 or older. Sources: 1950—*Statistical Abstract, 1954*, table 137. 1980—*1980 Census of Population, General Social and Economic Characteristics*, United States Summary, report PC80-1-C1, table 239. (+)

Table 5.1 (Continued)

HI-SCHOOL%	1950–80 change in high school graduates 25 and older as a percentage of all persons 25 or older. Sources: Same as for MD-YRS-SCHOOL. (+)
COLLEGE%	1950–80 change in college graduates as a percentage of all persons 25 or older. Sources: Same as for MD-YRS-SCHOOL. (+)
	WORKER-NONWORKER MIX
POP%EMPLOYED	1950–87 change in employment/population ratio, or total employed persons as a percentage of total population. Sources: 1950 employment and population—*County and City Data Book, 1956*, table 1. 1987 E/P ratio—precalculated in *Statistical Abstract, 1989*, table 623. (+)
POP%LABOR	1950–87 change in labor force participation rate, or total labor force as a percentage of total population. Sources: 1950—*1950 Census of Population*, v. 2, *Characteristics of the Population, United States Summary*, table 72. 1987—precalculated in *Statistical Abstract, 1989*, table 622. (+)
UNEMPLOYED%	1950–87 change in civilian unemployment rate, or total unemployment as a percentage of total labor force. 1950 total unemployment—1950 labor force (POP%LABOR source) minus 1950 employment (POP%EMPLOYED source). Unemployment was divided by labor force to get the 1950 unemployment rate. 1987 unemployment rate—*Statistical Abstract, 1989*, table 653. (-)
	HIGH-INCOME AND LOW-INCOME GROUPS
POP%BLACK	1950–88 change in black percentage of total population. Sources: 1950—*Statistical Abstract, 1954*, table 28. 1988—*Statistical Abstract, 1989*, table 31. (-)
POP%SPAIN	1970–80 change in Hispanic percentage of total population. Sources: 1970—*1970 Census of Population General Social and Economic Characteristics, United States Summary*, report PC(1)-C1, tables 146 and 147. 1980—*County and City Data Book, 1983*, table A. (-)
POP%INDIAN	1950–1980 change in Indian percentage of total population. Sources: 1950—*County and City Data Book 1956*, table 1, nonwhite percentage minus 1950 black percentage from POP%BLACK source. 1980—*County and City Data Book, 1983*, table A. (-)

Table 5.1 (Continued)

POP%WID-DIVOR	1950–80 change in widowed and divorced females as a percentage of total population. 1950—*1960 Census of Population*, v. 1, *General Population Characteristics*, chapter 5, table 105. 1980—*1980 Census of Population, General Population Characteristics, United States Summary*, report no. PC80-1-B1, table 651. Note: The reference to a *1960* census volume for *1950* data is correct. The 1950 data, not published with the 1950 census, were computed and published with the 1960 census to permit 1950–60 comparisons. (-)
OLD-INMIGRATN	1975–80 in-migration flows for persons 60 and older. Source: Special tabulations of 1980 census data by the Center for Social Research on Aging: University of Miami, Coral Gables, Florida. (+)
AMENITIES	Dummy variable valued at one for states where at least one county had both (*a*) a 1970 population percentage of 15 percent or more for persons 65 and older *and* (*b*) a 1960–70 net migration rate of +7 percent or more. The variable identifies preferred retirement states, where the retiree attraction is presumed to be natural amenities. The 19 states valued at one are ME, NH, NJ, PA, MI, WI, MN, MO, TN, FL, AR, OK, TX, MT, OR, CO, NM, AZ, and CA. Source: Wheat (1986), appendix A. (+)
	INDUSTRY MIX
MFG/POP	1950–87 change in manufacturing employment (production workers) per capita. Sources: 1950—*Statistical Abstract, 1954*, table 241. 1987—*Statistical Abstract, 1989*, table 656. (+)
LOW-MFG/POP	1950–87 change in low-wage-industry manufacturing employment per capita. Low-wage manufacturing is defined as textiles, apparel, furniture, and leather (SICs 22, 23, 25, and 31). Sources: 1950 figures are interpolated from 1947 and 1954 employment reported in *1954 Census of Manufacturers*, v. 3, *Area Statistics*, state tables. 1987—BLS, *Employment, Hours, and Earnings, States and Areas, 1972–87*, report L2.3:2320, Vol. 5 (1989). (+ or -)
HI-MFG/POP	1950–87 change in high-wage-industry manufacturing employment per capita. High-wage manufacturing is defined as chemicals and petroleum (SICs 28 and 29). Sources: Same as for LOW-MFG/POP. (+)
HIMID-MFG/POP	1950–87 change in combined high-wage and mid-wage manufacturing employment (all but textiles, apparel, furniture, and leather) per capita. Sources: Same as for LOW-MFG/POP. (+)

Table 5.1 (Continued)

SERVICE/POP	1950–87 change in broadly defined service sector employment per capita. "Service" is wholesale and retail trade, finance-insurance-real estate, and "services" (narrowly defined), or SIC groups F, G, H, and I. Sources: 1950—*Statistical Abstract, 1954*, table 241. 1987—*Statistical Abstract, 1989*, table 656. (-)
HI-SERV/POP	1950–87 change in high-earnings service-industry and employment per capita. High-earnings service industry is defined as depository institutions, non-depository credit institutions, security and commodity brokers and dealers, insurance carriers, holding and other investment offices, business services, health services, and engineering-accounting-research-management services (SICs 60-63, 67, 73, 80, and 87). Sources: Same as LOW-MFG/POP. (+)
OIL-GAS/POP	1950–87 change in petroleum and natural gas extraction employment per capita. Sources: 1950—U.S. Department of Labor, Bureau of Labor Statistics (BLS), *Employment and Earnings, States and Areas, 1939–74*, report no. L2.3:1370-11 (1975). 1987—BLS, *Employment, Hours, and Earnings, States and Areas, 1972–87*, report L2.3:2320, v. 1-5 (1989). (+)
MINING/POP	1950–87 change in metal, coal, and mineral mining employment per capita. Sources: Same as OIL-GAS/POP. (+)
OIL&MINE/POP	1950–87 change in combined petroleum, natural gas, and mining employment per capita. Sources: 1950—*Statistical Abstract, 1954*, table 241. 1987—*Statistical Abstract, 1959*, table 656 (mining, not shown separately, is obtained by subtracting other industries from total employment—see footnote 1 of table 656). (+)
PETROLBRL/POP	1950–87 change in crude petroleum production (millions of barrels) per capita. Sources: 1950—*Statistical Abstract, 1964*, table 1021 (1946–55 average used for 1950). 1987—*Statistical Abstract*, 1989, table 1189. (+)
	MISCELLANEOUS
TRANSCOSTDROP	Distance from Manufacturing Belt's Old Core, where distance is a proxy for both 1950 factory-to-market shipping costs and the subsequent (1950–87) reduction in those costs (reduction assumed to be proportional to 1950 costs). (Mileages are generally measured from Pittsburgh, but from a closer city for some northern New England and southern states.) Source: Wheat (1986), appendix A. (-)

Table 5.1 (Continued)

WAGE-RATE	1950–87 change in average hourly earnings in manufacturing (production workers). Sources: 1950—*Statistical Abstract, 1954*, table 248. 1987—*Statistical Abstract, 1989*, table 1267. (+)
ENERGY74/POP	1974 employment in petroleum and gas extraction and in petroleum refining, combined (not a change variable). Source: BLS, *Employment, Hours, and Earnings, States and Areas, 1972–87*, report 62.3:2320, v. 1-5 (1989). (+)
ENERGY82/POP	1982 employment in petroleum and gas extraction and in petroleum refining, combined (not a change variable). Source: Same as ENERGY74/POP. (-)
ENERGY-CRISIS	Dummy valued at one for 11 states where earnings from oil-gas extraction and coal mining accounted for at least 3 percent of the state's total 1981 earnings—states where energy price declines after 1981 had the most adverse effect. The states are WV, KY, LA, ND, MT, WY, CO, UT, OK, TX, NM. Source: Cletus C. Coughlin and Thomas B. Mandelbaum, "Why Have State Per Capita Incomes Diverged Recently?" *Federal Reserve Bank of St. Louis Review 70* (1988), table 4. (-)
FARM-CRISIS	Dummy valued at one for 12 states where earnings from farming accounted for at least 4 percent of the state's total 1981 earnings—states where farm price declines after 1981 had the most adverse effect. The states are VT, KY, WI, MN, IA, AK, ND, SD NB, KS, MT, and ID. Source: Coughlin and Mandelbaum (above), table 5. (-)
UNEMPLOY%1987	1987 civilian unemployment rate (proxy for short-run developments occurring in final decade of 1950–87 period). Source: *Statistical Abstract, 1989*, table 623. (-)
	REGIONAL DUMMIES AND INTERACTION VARIABLES
SLAVE	Dummy variable valued at one for the 15 slave states: DE, MD, VA, NC, SC, GA, FL, KY, TN, AL, MS, MO, AR, LA, and TX. (+)
NONSLAVE	Dummy variable valued at one for the 33 nonslave states. Source: Computed as -1 x (SLAVE - 1). (-)
SLxPOP%METRO	SLAVE x POP%METRO (+)
NSxPOP%METRO	NONSLAVE x POP%METRO (+)
SLxPOP%URBAN	SLAVE x POP%URBAN (+)
NSxPOP%URBAN	NONSLAVE x POP%URBAN (+)

Table 5.1 (Continued)

SLxPOP%FARM	SLAVE x POP%FARM (-)
NSxPOP%FARM	NONSLAVE x POP%FARM (-)
SLxPOP%TENANT	SLAVE x POP%TENANTFARM (-)
NSxPOP%TENANT	NONSLAVE x POP%TENANTFARM (-)
SLxFARM%xAC	SLAVE x FARM%xACRES (-)
NSxFARM%xAC	NONSLAVE x FARM%xACRES (-)
SLxTENANT%xAC	SLAVE x TENANT%xACRES (-)
NSxTENANT%xAC	NONSLAVE x TENANT%xACRES (-)
SLxPOP%BLACK	SLAVE x POP%BLACK (-)
NSxPOP%BLACK	NONSLAVE x POP%BLACK (-)

Note: Each of the change variables (all variables except AMENITIES, OLD-INMIGRATN, TRANSCOSTDROP, ENERGY74/POP, ENERGY82/POP, ENERGY-CRISIS, FARM-CRISIS, SLAVE, and NONSLAVE) has two versions, both of which use the same label. One version, used with the Percentage Change model, computes change the way INCOME/POP:% does. The other version, used with the State/Nation model, computes change the way INCOME/POP:S/N does.

Table 5.2
Simple Correlations: Percentage Change Model

	POP%METRO	POP%FARM	POP%TENANT FARM	SLxPOP% URBAN	SLxPOP% METRO	SLxPOP% TENANT
INCOME/POP%	.5159**	-.5317**	-.5198**	.6899**	.6139**	-.5839**
POP%METRO	1.0000	-.1701	-.4666**	.5162**	.6400**	-.3131*
POP%FARM	-.1701	1.0000	.3352*	-.4236**	-.3602*	.4057**
POP%TENANTFARM	-.4666**	.3352*	1.0000	-.6753**	-.5677**	.6981**
SLxPOP%URBAN	.5162**	-.4236**	-.6753**	1.0000	.9219**	-.9006**
SLxPOP%METRO	.6400**	-.3602**	-.5677**	.9219**	1.0000	-.7249**
SLxPOP%TENANT	-.3131*	.4057**	.6981**	-.9006**	-.7249**	1.0000
ACRES/FARM	.4764**	-.0283	-.4859**	.6612**	.7792**	-.4800**
MD-YRS-SCHOOL	.3779**	-.1582	-.6201**	.6826**	.5748**	-.6780**
HI-SCHOOL%	.3318*	-.3000*	-.5984**	.7441**	.6139**	-.7112**
COLLEGE%	.3793**	.0552	-.1989	.4555**	.4523**	-.3698**
WAGE-RATE	.3757**	-.2056	-.4555**	.5054**	.4369**	-.5593**
SLAVE	.2950*	-.3952**	-.6871**	.8844**	.7041**	-.9984**
POP%BLACK	-.4432**	.4158**	.5758**	-.8472**	-.8201**	.7244**
SLxPOP%BLACK	-.3888**	.2194	.1885	-.3844**	-.4751**	.0891
POP%EMPLOYED	.0033	-.5215**	-.0148	.0259	-.0858	-.0960
TRANSCOSTDROP	-.0611	-.0627	.1149	-.1046	-.0398	.1469
MFG/POP	.3805**	.2278	-.2329	.3032*	.3933**	-.1276
HIMID-MFG/POP	.4342**	.0082	-.4187**	.5192**	.5148**	-.3952**
UNEMPLOY%1987	.0769	-.1025	-.2826*	.2472*	.3068*	-.1702
UNEMPLOYED%	.2264	.3338*	-.3851**	.3363*	.3480**	-.2789*
AMENITIES	.2448*	-.0954	.0323	-.1130	-.0608	.0950
POP%WID-DIVOR	.0307	.1288	-.1028	.0992	.1343	.0005
OIL-GAS/POP	-.0142	-.0942	-.0489	-.0497	-.0522	.0397
OLD-INMIGRATN	.1550	-.2779*	-.1071	.0355	.0633	-.0651

	ACRES/FARM	MD-YRS-SCHOOL	HI-SCHOOL%	COLLEGE%²	WAGE-RATE	SLAVE
INCOME/POP%	.2686**	.6188**	.5686**	.5687**	.5934**	.5672**
POP%METRO	.4764**	.3779**	.3318*	.3793**	.3757**	.2950*
POP%FARM	-.0283	-.1582	-.3000*	.0552	-.2056	-.3952**
POP%TENANTFARM	-.4859**	-.6201**	-.5984**	-.1989	-.4555**	-.6871**
SLxPOP%URBAN	.6612**	.6826**	.7441**	.4555**	.5054**	.8844**
SLXPOP%METRO	.7792**	.5748**	.6139**	.4523**	.4369**	.7041**
SLxPOP%TENANT	-.4800**	-.6780**	-.7112**	-.3698**	-.5593**	-.9984**
ACRES/FARM	1.0000	.5441**	.5239**	.3800**	.4058**	.4539**
MD-YRS-SCHOOL	.5441**	1.0000	.8454**	.4751**	.5661**	.6636**
HI-SCHOOL%	.5239**	.8454**	1.0000	.3600**	.3936**	.6917**
COLLEGE%	.3800**	.4751**	.3600**	1.0000	.4638**	.3550**
WAGE-RATE	.4058**	.5661**	.3936**	.4638**	1.0000	.5514**
SLAVE	.4539**	.6636**	.6917**	.3550**	.5514**	1.0000
POP%BLACK	-.6320**	-.5990**	-.7366**	-.3853**	-.4177**	-.6974**
SLxPOP%BLACK	-.3783**	-.1943	-.2999*	-.1772	-.1136	-.0489
POP%EMPLOYED	-.3681**	-.1094	-.0200	-.0322	.2024	.0972
TRANSCOSTDROP	-.0232	-.4519**	-.1250	-.2436*	-.4689**	-.1520
MFG/POP	.4109**	.1867	.2287	.3246*	-.0353	-.1155
HIMID-MFG/POP	.4115**	.4200**	.5420**	.3080*	.1159	.3790**
UNEMPLOY%1987	.5056**	.2013	.2200	.0139	-.0330	.1614
UNEMPLOYED%	.5979**	.3882**	.3125*	.1800	.2181	.2732*
AMENITIES	-.2613*	-.0901	-.1576	.0191	.0075	-.0862
POP%WID-DIVOR	.1560	.3588**	.2640*	.1248	-.2868*	-.0095
OIL-GAS/POP	-.1066	-.1042	.0100	.0411	-.0793	-.0408
OLD-INMIGRATN	-.2945*	-.2180	-.0985	-.2077	-.1088	.0696

Table 5.2 (Continued)

	POP%BLACK	SLxPOP% BLACK	POP% EMPLOYED	TRANSCOST-DROP	MFG/POP	HIMID-MFG/POP
INCOME/POP%	-.6089**	-.3580**	.3999**	-.3587**	.1514	.4071**
POP%METRO	-.4432**	-.3888**	.0033	-.0611	.3805**	.4342**
POP%FARM	.4158**	.2194	-.5215**	-.0627	.2278	.0082
POP%TENANTFARM	.5758**	.1885	-.0148	.1149	-.2329	-.4187**
SLxPOP%URBAN	-.8472**	-.3844**	.0259	-.1046	.3032*	.5192**
SLXPOP%METRO	-.8201**	-.4751**	-.0858	-.0398	.3933**	.5148**
SLxPOP%TENANT	.7244**	.0891	-.0960	.1469	-.1276	-.3952**
ACRES/FARM	-.6320**	-.3783**	-.3681*	-.0232	.4109**	.4115**
MD-YRS-SCHOOL	-.5990**	-.1943	-.1094	-.4519**	.1867	.4200**
HI-SCHOOL%	-.7366**	-.2999*	-.0200	-.1250	.2287	.5420**
COLLEGE%	-.3853**	-.1772	-.0322	-.2436*	.3246*	.3080*
WAGE-RATE	-.4177**	-.1136	.2024	-.4689**	-.0353	.1159
SLAVE	-.6974**	-.0489	.0972	-.1520	.1155	.3790**
POP%BLACK	1.0000	.5830**	.0781	-.0886	-.3906**	-.6140**
SLxPOP%BLACK	.5830**	1.0000	.0999	-.1449	-.3096*	-.3847**
POP%EMPLOYED	.0781	.0999	1.0000	-.1282	-.2177	-.0184
TRANSCOSTDROP	-.0886	-.1449	-.1282	1.0000	.3258*	.1285
MFG/POP	-.3906**	-.3096*	-.2177	.3258*	1.0000	.8285**
HIMID-MFG/POP	-.6140**	-.3847**	-.0184	.1285	.8285**	1.0000
UNEMPLOY%1987	-.3124*	-.2186	-.5869**	.3036*	.1631	-.0558
UNEMPLOYED%	-.3258*	-.1451	-.6283**	-.0520	.3267*	.1944
AMENITIES	.1259	-.0336	.0582	.1770	.0017	-.1030
POP%WID-DIVOR	-.1186	-.1435	-.3232*	-.1209	.2763*	.2328
OIL-GAS/POP	.0123	.0181	.1083	.0382	.0201	.0108
OLD-INMIGRATN	-.1522	-.2226	.2542*	.4180**	.2274	.2155

	UNEMPLOY %1987	UNEMPLOYED%	AMENITIES	POP%WID-DIVOR	OIL-GAS/POP	OLD-INMIGRATN
INCOME/POP%	-.1492	-.0548	.0497	-.0124	.0143	.0589
POP%METRO	.0769	.2264	.2448*	.0307	-.0142	.1550
POP%FARM	-.1025	.3338*	-.0954	.1288	-.0942	-.2779*
POP%TENANTFARM	-.2826*	-.3851**	.0323	-.1028	-.0489	-.1071
SLxPOP%URBAN	.2472*	.3363*	-.1130	.0992	-.0497	.0355
SLXPOP%METRO	.3068*	.3480*	-.0608	.1343	-.0522	.0633
SLxPOP%TENANT	-.1702	-.2789*	.0950	.0005	.0397	-.0651
ACRES/FARM	.5056**	.5979**	-.2613*	.1560	-.1066	-.2945*
MD-YRS-SCHOOL	.2013	.3882**	-.0901	.3588**	-.1042	-.2180
HI-SCHOOL%	.2200	.3125*	-.1576	.2640*	.0100	-.0985
COLLEGE%	.0139	.1800	.0191	.1248	.0411	-.2077
WAGE-RATE	-.0330	.2181	.0075	-.2868*	-.0793	-.1088
SLAVE	.1614	.2732*	-.0862	-.0095	-.0408	.0696
POP%BLACK	-.3124*	-.3258*	.1259	-.1186	.0123	-.1522
SLxPOP%BLACK	-.2186	-.1451	-.0336	-.1435	.0181	-.2226
POP%EMPLOYED	-.5869**	-.6283**	.0582	-.3232*	.1083	.2542*
TRANSCOSTDROP	.3036*	-.0520	.1770	-.1209	.0382	.4180**
MFG/POP	.1631	.3267*	.0017	.2763*	.0201	.2274
HIMID-MFG/POP	-.0558	.1944	-.1030	.2328	.0108	.2155
UNEMPLOY%1987	1.0000	.6836**	.0498	.2430*	.0890	-.1018
UNEMPLOYED%	.6836**	1.0000	-.1503	.2299	.0621	-.2608*
AMENITIES	.0498	-.1503	1.0000**	-.0926	.1258	.3063*
POP%WID-DIVOR	.2430*	.2299	-.0926	1.0000	-.1735	-.0224
OIL-GAS/POP	.0890	.0621	.1258	-.1735	1.0000	.0289
OLD-INMIGRATN	-.1018	-.2608*	.3063*	-.0224	.0289	1.0000

* Significant at the .05 level (one-tail test); ** significant at the .01 level (one-tail test).

Correlations Involving the Dependent Variable

Variables representing several factors have significant, and sometimes strong, *r*'s with the dependent variable.[2] The strongest factors, ranked here by the *r* of the strongest variable measuring each factor, are (1) farm–urban mix, (2) education, (3) racial mix, (4) hourly wage rate, (5) industry mix, (6) employment, or worker-nonworker mix, (7) transport cost reductions on goods shipped from the Manufacturing Belt to western states, and (8) farm size. Remember that the variables describing these factors measure *change*, not levels (except that the transportation variable uses level as a proxy for change). Although we have heretofore used the term *rural*-urban mix to refer to the first factor, we will hereafter say *farm*-urban mix: the rural nonfarm population sector ultimately proves to be unimportant, so we emphasize what is important—the shift from *farming* (not all rural sectors) to urban places, occupations, and industries.

The strongest variable overall is SLxPOP%URBAN (+.69), or the 1950–87 percentage increase in the urban population percentage in former slave slaves. (The nonslave states are "zeroed out" by the slave state dummy, whose label is condensed to SL in the interaction variable's label.) Other farm–urban variables with high simple *r*'s include SLxPOP%METRO (+.61), SLxPOP%TENANT (-.58), POP%FARM (-.53), POP%TENANTFARM (-.52), and POP%METRO (+.52). These findings suggest three things: (1) big decreases in the farm and tenant farm (includes sharecropper) population percentages go with big increases in per-capita income, (2) big increases in the urban and metropolitan population percentages go with big increases in per-capita income, and (3) farm–urban mix changes have the most effect in the South, that is, in the former slave states (dummied in by the SLAVE multiplier). The South, of course, had by far the most profound changes in farm–urban mix; the demise of sharecropping is the reason. Moreover, the farm earnings level was a much lower percentage of the urban earnings level in the South than elsewhere: farm-to-urban population shifts of a given magnitude had more potential for raising income in the South than in other regions.

Education variables rank close behind, and sometimes outperform, the farm–urban variables. MD-YRS-SCHOOL (+.62) is the strongest educational variable and outperforms all other variables except SLxPOP%URBAN. Since the SLAVE multiplier (SL) gives SLxPOP% URBAN certain proxy characteristics (until other things are controlled), SLxPOP%URBAN to some extent represents education and racial mix in addition to farm–urban mix. If we therefore compare the best *noninteraction* variable describing farm–urban mix with the best education variable, MD-YRS-SCHOOL outperforms POP%FARM by +.62 to -.53. HI-SCHOOL% (+.57) and COLLEGE%^{2} (+.57) also outperform POP%FARM. (The college variable is squared to capture curvilinear correlation; COLLEGE%—unsquared—reads +.54.) These education *r*'s signify that big increases in educational attainment are associated with big increases in per-capita income.

Racial mix, best represented by POP%BLACK (-.61), ranks third by the simple *r* test. To anticipate, this variable gets short shrift in the regression equations, because it is cumulatively duplicated by farm–urban, education, and

employment variables. All four factors—racial mix, farm-urban mix, education, and employment—had their biggest 1950–87 changes in the South; the collapse of sharecropping and the attendant migration of black former sharecroppers to southern cities and nonsouthern regions link the four factors. When we get to the regression analysis, the duplication of racial mix by the other three factors will give a false impression that racial mix is a relatively weak influence. We should not forget, therefore, that the simple *r*'s tell a different story: POP%BLACK is .01 away from being tied with MD-YRS-SCHOOL for the honor of having the highest simple *r* for a noninteraction variable.

The fourth-ranked (by simple *r*) factor, hourly wage rate, is represented by WAGE-RATE (+.59). Note that WAGE-RATE is only .03 behind MD-YRS-SCHOOL and outperforms the other two educational variables. Although WAGE-RATE specifically measures *manufacturing* (production worker) wage increases, it serves as proxy for the general level of the state's wage structure. Using a specific sector's (manufacturing's) wage rate provides some control over state-to-state variations in industry mix; WAGE-RATE thus avoids some of the industry mix bias that would weaken a more general wage variable. The positive sign means that the biggest per-capita income increases have tended to occur in the states with the biggest hourly wage increases. (A +.55 intercorrelation between WAGE-RATE and SLAVE tells us that the biggest wage increases have been in the South, where wages are catching up with nonsouthern wages.)

The fifth-ranked factor is industry mix. This factor could be described more narrowly as manufacturing-nonmanufacturing mix: service and resource industry variables were tested in addition to manufacturing variables, but only the manufacturing variables were significant in correlation or regression tests. The best manufacturing variable by correlation test was HIMID-MFG/POP, or per-capita employment in the combined high-wage and mid-wage manufacturing industries. The high-wage industries are chemicals and petroleum refining; the mid-wage industries are everything else except textiles, apparel, leather, and furniture. HIMID-MFG/POP's *r* of +.41 is comfortably above the 1 percent level (.33). But LOW-MFG/POP (textiles, apparel, leather, and furniture) has a thoroughly insignificant *r* of -.02; and MFG/POP, which embraces all manufacturing, has an *r* of +.15, well short of the 5 percent level (.24). The correlation evidence suggests, in short, that growth in only the non-low-wage manufacturing industries contributes to income gains. When we get to the regression equations, however, we shall see that MFG/POP (*all* manufacturing) occasionally outperforms HIMID-MFG/POP.

The sixth strongest factor is employment, or the worker-nonworker mix of the population. The main variable is POP%EMPLOYED (+.40), which amounts to the same thing as the employment/population ratio (E/P). POP%EMPLOYED's simple *r* is deceptively low: variation in more powerful influences partly hides the effect of variation in employment. But controlling the strongest farm-urban variable, SLxPOP%URBAN, changes POP%EMPLOYED's simple *r* of +.40 to a partial *r* of +.53. Alternatively, controlling the strongest education variable, MD-YRS-SCHOOL, sends POP%EMPLOYED to +.60. Controlling both MD-YRS-SCHOOL and COLLEGE%^{2} sends POP%EMPLOYED to +.64. And controlling three variables—MD-YRS-SCHOOL, COLLEGE%^{2}, and *either*

POP%BLACK or SLxPOP%METRO—gives POP%EMPLOYED a partial *r* of +.70. These *r*'s reveal a strong latent relationship between income and employment.

Seventh rank goes to the western transportation cost reduction factor. TRANSCOSTDROP (-.36) describes 1950–87 reductions in the cost of shipping manufactured goods from the Manufacturing Belt to external states. The western states have the biggest transport cost reductions. The *r* of -.36 exceeds the 1 percent significance level. High values for TRANSCOSTDROP represent big decreases in the cost of shipping goods from the Manufacturing Belt to external states. TRANSCOSTDROP's negative sign therefore means that the highest shipping cost reductions are associated with the lowest per-capita income increases, that is, with *relative* decreases. Lower delivered prices on eastern manufactured goods sold in the West seem to have lowered the West's price-wage structure and thereby lowered the West's per-capita income relative to the East's.

The only other factor offering a significant *r* is farm size. ACRES/FARM (+.27), or increase in average acres per farm, is significant at only about the 3 percent level. This variable assesses the effect of farm consolidation on income. Bigger farms should mean more income per farmer. The effect is visible but weak. An intercorrelation of +.45 between ACRES/FARM and SLAVE says that the biggest increases in average acreage were in the South. And this finding, viewed in the light of ACRES/FARM's relative weakness, suggests that ACRES/FARM could be acting as a proxy for the combination of influences that gave fast income growth to the South. Some other intercorrelations amplify this suggestion: ACRES/FARM has *r*'s of +.78 with SLxPOP%METRO, +.52 with POP%URBAN, -.63 with POP%BLACK, and +.54 with MD-YRS-SCHOOL.

Intercorrelations Involving SLAVE

Some of the most illuminating of table 5.2's intercorrelations are those for SLAVE, the dummy valued at one for former slave states and at zero elsewhere. SLAVE's *r*'s show which variables tend to have their biggest increases or decreases in the South. These are the variables that should contribute the most to an explanation of why the South's income closed in on that of the nonsouthern regions. SLAVE has high *r*'s with variables measuring changes in (*a*) racial mix, (*b*) farm-urban mix, (*c*) educational attainment, (*d*) wage rates, (*e*) farm size, and (*f*) industry mix. Variables representing these factors have the following *r*'s with SLAVE:

-.70	POP%BLACK	+.66	MD-YRS-SCHOOL
-.69	POP%TENANTFARM	+.36	COLLEGE%^{2}
-.40	POP%FARM	+.55	WAGE-RATE
+.38	POP%URBAN	+.45	ACRES/FARM
+.69	HI-SCHOOL%	+.38	HIMID-MFG/POP

SLAVE's *r*'s of -.70 for POP%BLACK and -.69 for POP%TENANTFARM reflect the collapse of sharecropping and the attendant out-migration from the

South of millions of blacks. The two *r*'s reiterate two points: (1) the biggest decreases in black population percentage (i.e., the most negative values of POP%BLACK) were in southern states, where SLAVE's value is highest (one); (2) the biggest decreases in tenant farmer and sharecropper population (i.e., the most negative values of POP%TENANTFARM) were in the southern states. The out-migration of undereducated blacks—and poor whites—from the South contributed to the rise in southern educational attainment. This development registers in SLAVE's *r*'s of +.69 with HI-SCHOOL% and +.66 with MD-YRS-SCHOOL. Sharecropping's virtual disappearance led to gradual evaporation of the South's surplus labor, whose source was agriculture. The result was rising wages in the South; the *r* of +.55 for WAGE-RATE describes this. Finally, the plantations that had been subdivided into tenant-sharecropper tracts were recombined into single farms. The resulting increases in average farm acreage in the South (hence in average income per farmer) explain the +.45 *r* relating ACRES/FARM to SLAVE.

The last of the above-listed SLAVE intercorrelations, the +.38 for HIMID-MFG/POP, is not explained by the collapse of sharecropping. Instead, it is explained by rapid manufacturing growth in the South—another source of income gains. Oddly, SLAVE has an *r* of only +.12 with MFG/POP, the variable that covers low-wage in addition to high-wage and mid-wage manufacturing. Growth in low-wage industries—mainly textiles and apparel—has been a widely publicized element of southeastern growth. But textile employment peaked in 1968, and apparel employment peaked in 1973. Since then, employment in both industries has been declining. The declines are due partly to improved productivity (new technology) and partly to production shifts to less developed countries. The western part of the South always did emphasize high-wage and mid-wage manufacturing.

The ten SLAVE intercorrelations listed cover all but one of the factors theorized to have contributed to the South's extraordinary income gains. The one exception is employment (E/P). POP%EMPLOYED does have a right-sign *r* of +.10 with SLAVE, but that *r* is not even close to the 5 percent significance level (.24). Three things explain this weakness. First, the very biggest E/P gains were in the West; the West had lots of in-migration by working-age males—a high labor force participation group— from agricultural states. Second, the Plains was only slightly behind the South in percentage increase in E/P. Farm population decline was qualitatively different in the Plains: the farms that disappeared or were consolidated were owner occupied, so the wives were not in the labor force. Urbanization in the Plains thus replaced a low E/P group (farmers) with a high E/P group (urban residents). Third—this really is an extension of the second reason—sharecropper decline in the South involved decline in a moderately high participation group: the women on sharecropper farms were field hands and thus boosted E/P.[3] Hence *part* of the South's farm decline (the sharecropper-tenant part) saw one high E/P group (sharecroppers) being replaced by another (urban residents). The South's E/P gains resulted not so much from sharecropping's decline as from owner-occupied farm decline, improved education, and reduced employment discrimination against blacks.

Other Intercorrelations

The most useful of table 5.2's other intercorrelations involve either relationships (*a*) between variables describing the same factor or (*b*) among factors that have a southern emphasis. We begin with relationships among the farm-urban mix variables. The two farm variables—POP%FARM and POP%TENANTFARM—have an *r* of only +.34. This *r*'s relatively low magnitude reminds us that the southern states had big declines in tenant-sharecropper farming, whereas the nonsouthern state farm losses emphasized owner-occupied farms. The two urban variables —POP%URBAN and POP%METRO—also had just a moderate *r*: +.32. This *r*'s failure to be higher results from another regional contrast. The South is heavily rural, so most of its urban growth was in nonmetropolitan areas; but urban growth in the North favored metropolitan areas.

Because the association between farm decline and urban growth was strongest in the South, POP%TENANTFARM produces stronger intercorrelations than POP%FARM does. POP%TENANTFARM has *r*'s of -.32 with POP%URBAN, -.68 with SLxPOP%URBAN, -.47 with POP%METRO, and -.57 with SLxPOP%METRO. These *r*'s show that farm decline and urban growth are something like two sides of a coin; the effect of farm decline can register to a considerable extent in a variable that measures urban growth, and vice versa.

For education, an extremely high *r* of +.85 between MD-YRS-SCHOOL and HI-SCHOOL% explains why HI-SCHOOL% enters no equations. MD-YRS-SCHOOL is the stronger of the two and, once it enters an equation, duplicates HI-SCHOOL% to such an extent that the latter becomes impotent. COLLEGE% has somewhat weaker *r*'s of +.48 with MD-YRS-SCHOOL and +.36 with HI-SCHOOL%. The former *r* leaves COLLEGE% strong enough to enter many equations even when MD-YRS-SCHOOL is controlled. The very best equations obtainable with three or more variables always include both MD-YRS-SCHOOL and COLLEGE%^{2}.

Turning to relationships among factors that have a southern emphasis, we are naturally curious about the connections between farm-urban mix and education: these ultimately prove to be the two strongest factors. The best education variable, MD-YRS-SCHOOL, has *r*'s of +.35 with POP%URBAN, +.54 with POP%METRO, +.57 with SLxPOP%METRO, and +.68 with SLxPOP%URBAN. The SLAVE dummy multipliers in the last two variables include elements of education; this explains why the last two *r*'s are especially high. MD-YRS-SCHOOL also has *r*'s of -.16 with POP%FARM and -.62 with POP%TENANT farm. The remarkable contrast between these two *r*'s reiterates a now-familiar point: the biggest educational gains were in the South, where *tenant-sharecropper* farming declined precipitously; whereas the smallest educational gains were in the North, where declines in *conventional* farming were relatively strong.

POP%BLACK has a powerful simple *r* of -.61; the only stronger noninteraction variable is MD-YRS-SCHOOL, +.62. Yet POP%BLACK will perform poorly when we get to the regression equations. Why? Intercorrelations with variables representing education and farm-urban mix (the two strongest

factors) provide the answer. POP%BLACK has *r*'s of -.74 with HI-SCHOOL%, -.60 with MD-YRS-SCHOOL, -.39 with COLLEGE%^{2}, -.85 with SLxPOP%URBAN, +.74 with SLxPOP%TENANT, +.58 with POP%TENANTFARM, -.57 with POP%URBAN, and +.42 with POP%FARM. Several of these intercorrelated variables dominate the regression equations. The best two-variable and three-variable equations use farm urban mix and education variables exclusively; the best six-variable equation has five such variables, including SLxPOP%URBAN (-.85 *r* with POP%BLACK) and MD-YRS-SCHOOL (-.60). The cumulative effect of all this duplication on POP%BLACK is devastating. Simply holding SLxPOP%URBAN constant changes POP%BLACK's simple *r* of -.61 to a partial *r* of -.09. (That is another way of saying that the South's black population decline expressed itself as a rise in urbanization.) The wonder is that, in one equation, POP%BLACK regains significance despite the control of not only SLxPOP%URBAN but POP%TENANT and COLLEGE%^{2}.

Another variable that fades in the regression equations is WAGE-RATE. Its simple *r* of +.59 is right up there with those of +.62 for MD-YRS-SCHOOL and -.61 for POP%BLACK. But WAGE-RATE suffers from high intercorrelations with variables that prove to be more powerful. The troublesome intercorrelations include ones of +.57 with MD-YRS-SCHOOL, +.46 with COLLEGE%^{2}, +.51 with SLxPOP%URBAN, and -.46 with POP%TENANT. The other variables collectively duplicate the explanatory power of WAGE-RATE. Controlling MD-YRS-SCHOOL alone sends WAGE-RATE from +.59 to +.38.

A final pair of intercorrelations provides a reprise for the contrast between tenant-sharecropper farming in the South and conventional farming elsewhere. (Bear in mind that the South also had owner-occupied farms; they just weren't as severely affected by farm population's decline.) POP%EMPLOYED, or E/P (the employment/population ratio), has contrasting *r*'s of -.01 with POP%TENANTFARM and -.52 with POP%FARM. Although POP%FARM includes tenant sharecropper farms, most of the nation's farm population was on owner-occupied farms. On these farms, the farm wives did not work in the fields; they were *not* part of the E in E/P. But on tenant-sharecropper farms, the women were field hands. Because of this contrast, population shifts from conventional farms to urban areas have considerable effect on E/P: nonworking farm women tend to become working urban women. When population shifts from tenant-sharecropper farms to urban areas, though, working women remain working women; E/P does not climb. Hence the sharecropper variable, POP%TENANTFARM, shows no association between farm population change and E/P change.

The preceding analysis might seem to belie our earlier point that the South had relatively big increases in E/P. But there is no real contradiction. As stated in chapter 4, the biggest 1950-87 E/P increases were those of 74 percent in the West, 72 percent in the Deep South, and 71 percent in the Peripheral South. In contrast, both halves of the Manufacturing Belt –the Northeast and the Great Lakes—had E/P increases of only 56 percent; the Plains had a 70 percent increase. Some of the South's E/P increase resulted from the population decline on owner-occupied farms. The rest is largely attributable to educational gains and

to reductions in employment discrimination. The point we are now making is that, whereas the South had declines in both types of farming, only the decline in conventional farming (owner-occupied farms) had much effect on E/P.

State/Nation Model

Table 5.3 is a correlation matrix for the State/Nation model; the variables are basically the same as those in table 5.2 (Percentage Change model). The State/Nation model's dependent variable is the 1950–87 absolute increase in the ratio of state per-capita income to national per-capita income. Table 5.3's *r*'s are generally quite similar to the comparable *r*'s in table 5.2 but tend to be slightly lower. In both models the two strongest *variables* are SLxPOP%URBAN (strongest) and MD-YRS-SCHOOL (second-strongest). And in both models the three strongest *factors*, in one-two-three order, are (1) farm-urban mix, (2) educational attainment, and (3) the hourly wage rate. Variables that have high intercorrelations with SLAVE in the Percentage Change model usually have high intercorrelations in the State/Nation model too. The *r*'s for SLAVE, however, display less similarity between models than those for the other independent variables.

Because the two models provide generally similar simple *r*'s, the State/Nation model does not require detailed analysis. As a substitute for detailed analysis, table 5.4 provides a summary comparison of the two most important sets of *r*'s from the models—the *r*'s involving the dependent variables and the *r*'s involving SLAVE. The stub column lists the 24 variables shown in tables 5.2 and 5.3. The next two columns give the simple *r*'s of those variables with the dependent variables. The last two columns give the intercorrelations between the stub column variables and SLAVE, first for the Percentage Change model and then for the State/Nation model.

Seventeen of table 5.4's variables have significant *r*'s (5 percent level or better) with the dependent variable in at least one model. Of these 17 variables, 13 are more significant in the Percentage Change model, 3 are more significant in the State/Nation model, and 1 (ACRES/FARM) is +.27 in both models. The Percentage Change model thus has about four times as many highest *r*'s. In addition, the Percentage Change model predicts best for the four factors whose variables have the highest simple *r*'s—farm-urban mix, educational attainment, hourly wages, and racial mix.

Table 5.3
Simple Correlations: State/Nation Model

	POP%METRO	POP%FARM	POP%TENANT FARM	SLxPOP% URBAN	SLxPOP% METRO	SLxPOP% TENANT
INCOME/POP%	.0273	-.4266**	-.4832**	.5899**	.5050**	-.3983**
POP%METRO	1.0000	.0760	.0417	.2271	.3348*	-.1746
POP%FARM	.0760	1.0000	.7525**	-.4850**	-.3152*	.4178**
POP%TENANTFARM	.0417	.7525**	1.0000	-.5668**	-.3657**	.6266**
SLxPOP%URBAN	.2271	-.4850**	-.5668**	1.0000	.7639**	-.7472**
SLXPOP%METRO	.3348*	-.3152*	-.3657**	.7639**	1.0000	-.5739**
SLxPOP%TENANT	-.1746	.4178**	.6266**	-.7472**	-.5739**	1.0000
ACRES/FARM	-.2240	.0317	-.1266	.2137	.0986	-.1752
MD-YRS-SCHOOL	.0580	-.1440	-.3405**	.5950**	.3805**	-.4845**
HI-SCHOOL%	.3109*	-.0358	-.1156	.4192**	.2599*	-.3263*
COLLEGE%	.1294	.1138	-.0024	.2941*	.1842	-.2738*
WAGE-RATE	-.0965	-.1808	-.2053	.3794**	.1575	-.2791*
SLAVE	.0600	-.4421**	-.5022**	.7868**	.3198*	-.5140**
POP%BLACK	-.2070	.4295**	.5121**	-.8605**	-.7061**	.7712**
SLxPOP%BLACK	-.2419*	.4663**	.5810**	-.8959**	-.7861**	.8330**
POP%EMPLOYED	-.0735	-.4195**	-.2476*	.0307	-.1026	.0658
TRANSCOSTDROP	.2626*	-.0273	.0701	-.0629	.0884	.0182
MFG/POP	.4335**	.0435	-.1397	.5335**	.4401**	-.4762**
HIMID-MFG/POP	.3621**	.0039	-.1482	.4523**	.3754**	-.3854**
UNEMPLOY%1987	.1259	-.2390	-.1770	.2508*	.2867*	-.2822*
UNEMPLOYED%	.1669	.0860	-.0382	.3471**	.2397	-.3273*
AMENITIES	.1051	-.1009	-.0244	-.0918	.0049	.1891
POP%WID-DIVOR	.0884	.1370	-.0131	.1381	.1521	-.1848
OIL-GAS/POP	-.1108	.0035	-.0760	.1006	.0565	-.1419
OLD-INMIGRATN	.3772**	-.1560	-.0583	.0759	.1231	.0817

Table 5.3 (Continued)

	ACRES/FARM	MD-YRS-SCHOOL	HI-SCHOOL%	COLLEGE%[2]	WAGE-RATE	SLAVE
INCOME/POP%	.2715*	.5661**	.3004*	.5132**	.5422**	.4651**
POP%METRO	-.2240	.0580	.3109*	.1294	-.0965	.0600
POP%FARM	.0317	-.1440	-.0358	.1138	-.1808	-.4421**
POP%TENANTFARM	-.1266	-.3405**	-.1156	-.0024	-.2053	-.5022**
SLxPOP%URBAN	.2137	.5950**	.4192**	.2941*	.3794**	.7868**
SLXPOP%METRO	.0986	.3805**	.2599*	.1842	.1575	.3198*
SLxPOP%TENANT	-.1752	-.4845**	-.3263*	-.2738*	-.2791*	-.5140**
ACRES/FARM	1.0000	.2942*	-.1089	.3072*	.3919**	.2461*
MD-YRS-SCHOOL	.2942*	1.0000	.6732**	.4115**	.5443**	.6267**
HI-SCHOOL%	-.1089	.6732**	1.0000	.4705**	.1692	.4569**
COLLEGE%	.3072*	.4115**	.4705**	1.0000	.3931**	.2401*
WAGE-RATE	.3919**	.5443**	.1692	.3931**	1.0000	.4865**
SLAVE	.2461*	.6267**	.4569**	.2401*	.4865**	1.0000
POP%BLACK	-.0968	-.5373**	-.4437**	-.2319	-.3154*	-.6974**
SLxPOP%BLACK	-.1907	-.5601**	-.4063**	-.2008	-.3595**	-.6691**
POP%EMPLOYED	-.0981	-.1121	-.0910	.0435	.2189	.1057
TRANSCOSTDROP	-.4370**	-.5130**	-.0326	-.2917*	-.5173**	-.1520
MFG/POP	-.0586	.2281	.3105*	.2128	.0704	.3730**
HIMID-MFG/POP	-.0254	.2159	.2583*	.2538*	.0862	.3228*
UNEMPLOY%1987	-.0939	.1630	.1821	-.1032	-.0731	.1614
UNEMPLOYED%	.1484	.4278**	.3497*	.0684	.1572	.3468**
AMENITIES	-.1979	-.0530	.0559	.0248	.0277	-.0862
POP%WID-DIVOR	-.0664	.4181**	.4286**	.0874	-.2505*	.0238
OIL-GAS/POP	.3511**	.2576*	.0375	.2084	.3497**	.1108
OLD-INMIGRATN	-.5440**	-.2105	.0558	-.2429*	-.1469	.0696

	POP%BLACK	SLxPOP% BLACK	POP% EMPLOYED	TRANSCOST-DROP	MFG/POP	HIMID-MFG/POP
INCOME/POP%	-.4868**	-.4996**	.4416**	-.3875**	.2116	.3790**
POP%METRO	-.2070	-.2419*	-.0735	.2626*	.4335**	.3621**
POP%FARM	.4295**	.4663**	-.4195**	-.0273	.0435	.0039
POP%TENANTFARM	.5121**	.5810**	-.2476*	.0701	-.1397	-.1482
SLxPOP%URBAN	-.8605**	-.8959**	.0307	-.0629	.5335**	.4523**
SLXPOP%METRO	-.7061**	-.7861**	-.1026	.0884	.4401**	.3754**
SLxPOP%TENANT	.7712**	.8330**	.0658	.0182	-.4762**	-.3854**
ACRES/FARM	-.0968	-.1907	-.0981	-.4370**	-.0586	-.0254
MD-YRS-SCHOOL	-.5373**	-.5601**	-.1121	-.5130**	.2281	.2159
HI-SCHOOL%	-.4437**	-.4063**	-.0910	-.0326	.3105*	.2583*
COLLEGE%	-.2319	-.2008	.0435	-.2917*	.2128	.2538*
WAGE-RATE	-.3154*	-.3595**	.2189	-.5173**	.0704	.0862
SLAVE	-.6974**	-.6691**	.1057	-.1520	.3730**	.3228*
POP%BLACK	1.0000	.9321**	.0705	-.0886	-.5911**	-.5573**
SLxPOP%BLACK	.9321**	1.0000	.0471	.0015	-.4921**	-.4479**
POP%EMPLOYED	.0705	.0471	1.0000**	-.1418	-.1102	.1372
TRANSCOSTDROP	-.0886	.0015	-.1418	1.0000	.3294*	.1736
MFG/POP	-.5911**	-.4921**	-.1102	.3294*	1.0000	.8267**
HIMID-MFG/POP	-.5573**	-.4479**	.1372	.1736	.8267**	1.0000
UNEMPLOY%1987	-.3124*	-.2786*	-.5822**	.3036*	.2301	-.0990
UNEMPLOYED%	-.3585**	-.3461**	-.6322**	-.0288	.4179**	.0779
AMENITIES	.1259	.1256	.0401	.1770	.0215	-.0137
POP%WID-DIVOR	-.1330	-.1205	-.2989*	-.1545	.0899	-.0029
OIL-GAS/POP	-.0830	-.1286	.1771	-.1098	.0529	.0747
OLD-INMIGRATN	-.1522	-.1221	.2406*	.4180**	.1747	.2229

Table 5.3 (Continued)

	UNEMPLOY %1987	UNEMPLOYED%	AMENITIES	POP%WID-DIVOR	OIL-GAS/POP	OLD-INMIGRATN
INCOME/POP%	-.2110	-.1276	.0863	-.0264	.2721*	.0461
POP%METRO	.1259	.1669	.1051	.0884	-.1108	.3772**
POP%FARM	-.2390	.0860	-.1009	.1370	.0035	-.1560
POP%TENANTFARM	-.1770	-.0382	-.0244	-.0131	-.0760	-.0583
SLxPOP%URBAN	.2508*	.3471*	-.0918	.1381	.1006	.0759
SLXPOP%METRO	.2867*	.2397*	.0049	.1521	.0565	.1231
SLxPOP%TENANT	-.2822*	-.3273*	.1891	-.1848	-.1419	.0817
ACRES/FARM	-.0939	.1484	-.1979	-.0664	.3511**	-.5440**
MD-YRS-SCHOOL	.1630	.4278**	-.0530	.4181**	.2576*	-.2105
HI-SCHOOL%	.1821	.3497**	.0559	.4286**	.0375	.0558
COLLEGE%	-.1032	.0684	.0248	.0874	.2084	-.2429*
WAGE-RATE	-.0731	.1572	.0277	-.2505*	.3497*	-.1469
SLAVE	.1614	.3468**	-.0862	.0238	.1108	.0696
POP%BLACK	-.3124*	-.3585**	.1259	-.1330	-.0830	-.1522
SLxPOP%BLACK	-.2786*	-.3461**	.1256	-.1205	-.1286	-.1221
POP%EMPLOYED	-.5822**	-.6322**	.0401	-.2989*	.1771	.2406*
TRANSCOSTDROP	.3036*	-.0288	.1770	-.1545	-.1098	.4180**
MFG/POP	.2301	.4179**	.0215	.0899	.0529	.1747
HIMID-MFG/POP	-.0990	.0779	-.0137	-.0029	.0747	.2229
UNEMPLOY%1987	1.0000	.7536**	.0498	.2360	-.0686	-.1018
UNEMPLOYED%	.7536**	1.0000	-.1489	.2698*	-.0700	-.2471*
AMENITIES	.0498	-.1489	1.0000**	-.0898	.0099	.3063*
POP%WID-DIVOR	.2360	.2698*	-.0898	1.0000	-.0030	-.0120
OIL-GAS/POP	-.0686	-.0700	.0099	-.0030	1.0000**	.0080
OLD-INMIGRATN	-.1018	-.2471*	.3063*	-.0120	.0080	1.0000

* Significant at the .05 level (one-tail test); ** significant at the .01 level (one-tail test).

Table 5.4
Summary Comparison of Simple Correlations from Two Models

Independent Variable	Correlations with Dependent Variable		Correlations with SLAVE	
	% Change Model	State/ Nation Model	% Change Model	State/ Nation Model
POP%METRO	+.52	+.03	+.30	+.06
POP%FARM	-.53	-.43	-.40	-.44
POP%TENANTFARM	-.52	-.48	-.69	-.50
SLxPOP%URBAN	+.69	+.59	+.88	+.79
SLxPOP%METRO	+.61	+.51	+.70	+.32
SLxPOP%TENANT	-.58	-.40	-.99	-.51
ACRES/FARM	+.27	+.27	+.45	+.25
MD-YRS-SCHOOL	+.62	+.57	+.66	+.63
HI-SCHOOL%	+.57	+.30	+.69	+.46
COLLEGE%^{2}	+.57	+.51	+.36	+.24
WAGE-RATE	+.59	+.54	+.55	+.49
SLAVE	+.57	+.47	+1.00	+1.00
POP%BLACK	-.61	-.49	-.70	-.70
SLxPOP%BLACK	-.36	-.50	-.05	-.67
POP%EMPLOYED	+.40	+.44	+.10	+.11
TRANSCOSTDROP	-.36	-.39	-.15	-.15
MFG/POP	+.15	+.21	+.12	+.37
HIMID-MFG/POP	+.41	+.38	+.38	+.32
OIL-GAS/POP	+.01	+.27	-.04	+.11
UNEMPLOY%1987	-.15	-.21	+.16	+.16
UNEMPLOYED%	-.05	-.13	+.27	+.35
POP%WID-DIVOR	-.01	-.03	-.01	+.02
AMENITY	+.05	+.09	-.09	-.09
OLD-INMIGRATN	+.06	+.05	+.07	+.07

NOTES

1. Leonard F. Wheat, "The Determinants of 1963–77 Regional Manufacturing Growth: Why the South and West Grow," *Journal of Regional Science* 26 (1986), 650–51.

2. Our analysis will often refer to *factors* and *variables*. Many people use these terms interchangeably, as though they were synonyms. Viewing *factor* and *variable* as synonyms will lead to misunderstanding. The terms carry important distinctions. A *factor* is something general or abstract, such as climate or intelligence. A factor is too abstract to have a numerical value, whereas *variables* representing the factor have such values. A variable, then, measures a factor. A variable is a concrete (not abstract), measurable feature having values that *vary* (differ) from case to case (states, persons, years, etc.). July maximum temperature, average rainfall, and heating degree days are variables that measure climate, a factor. Some equations in this study have two or three variables measuring one factor. Our best equation has three variables measuring farm-urban population mix, two measuring educational attainment, and two measuring employment.

3. Gavin Wright, *Old South, New South: Revolutions in the Southern Economy Since the Civil War* (New York: Basic Books, 1986), 94; and Nicholas Lemann, *The Promised Land: The Great Black Migration and How It Changed America* (New York: Knopf, 1991), 7–8, 29.

6

REGRESSION ANALYSIS

In this chapter we examine 17 equations for the Percentage Change model and 14 for the State/Nation model. The Percentage Change model uses INCOME/POP:% (1950–87 percentage increase in state per-capita income) as the dependent variable; the State/Nation model uses INCOME/POP:S/N (1950–87 percentage *point* increase in relative per-capita income, i.e., in state per-capita income expressed as a percentage of U.S. per-capita income). The equations range in length from two variables to ten. Most of the equations are so-called *Best-R equations*: equations whose variables yield the highest R attainable for a given number of right-sign variables or else (for each model's longest equation) for an unlimited number of significant, right-sign variables. Additional equations show either (*a*) the variables that yield the second-highest R for a given number of variables or (*b*) a featured variable that seldom or never appears in a Best-R equation but is nevertheless significant when certain overlapping influences are not fully controlled. A final equation for each model, called the *consolidated equation*, takes the variables from the longest Best-R equation and combines (consolidates) all variables measuring a particular factor (e.g., education) into a single *compound variable*. The compound variable registers the significance (t) of the multivariable factor. The consolidated equation has one compound variable for each factor that is represented by more than one variable in the longest equation. Variables representing one-variable factors are not consolidated but are retained in the consolidated equation in their original form.

DETAILS OF REGRESSION METHODOLOGY

Before presenting and discussing the equations, we will clarify the purposes and details of our regression methodology.

Purposes of Using Many Equations and the Best-*R* Criterion

The reason for using many equations and two dependent variables is to facilitate better judgments about the relative importance of different factors. Variables (which represent factors) behave differently in different control contexts (equations). Some variables not in the longest (R_{10}) equation are in shorter equations, and vice versa. A variable that is strongest (highest t) in one equation may be second or fourth strongest in another or not even in it. We look for consistent performance by a factor (e.g., education), although different variables may represent the factor in different equations. Similarly, one dependent variable may show a stronger relationship to a certain factor than does the other dependent variable. In short, by consulting many equations we get a well-rounded picture.

Factors whose variables are strong enough to enter the shorter equations are relatively powerful; factors represented only in the longer equations—perhaps just the longest one—are generally weaker than those that enter the shorter equations. As discussed below, however, these rules do not always apply, particularly when multicollinearity is present.

Using R to determine which variables enter particular equations makes the analysis more objective, hence more reliable. Research risks bias when equations are strictly prespecified—where the researcher chooses all of the variables entering an equation. Prespecification lets the researcher determine judgmentally what is and is not controlled; the researcher's preconceptions color the results. Of course, when using R to determine which variables enter an equation, the researcher must limit the pool of available variables to theoretically plausible ones. Otherwise, the researcher is "data mining" and risking spurious correlation, wherein a variable is statistically significant because its "significance" occurs by chance. All variables used in this study are backed by persuasive hypotheses. We formulated all hypotheses before their variables were chosen for testing.

When using R as the criterion for entering variables into an equation, one must eliminate insignificant and wrong-sign variables. Insignificant and wrong-sign variables are often intercorrelated with other variables in the equation. Even an insignificant variable can greatly reduce the t value of a significant variable if the two variables have a strong partial intercorrelation. We realize that significant, right-sign variables can also have intercorrelation (multicollinearity) problems, but that is no excuse for accepting the avoidable intercorrelation that arises from insignificant and wrong-sign variables.

Alternate Equations

The regression tables include some "alternate" equations. These are not prespecified, but they are not the best or second-best either. The Best-R equations (and sometimes even the second-best) do have one drawback: they can hide influences that overlap others that are represented in the equations. In other words, one or more variables in a Best-R equation may duplicate some other genuinely significant variable to a sufficient extent that the duplicated variable is

rendered too weak to enter the equation. To a degree, variables in the equation serve as proxies for overlapping ones left out. The false impression arises that the ones left out always represent invalid hypotheses. Among the influences that tend to be hidden by other influences are changes in (*a*) manufacturing employment per capita (duplicated by farm and urban population variables), (*b*) average hourly wage rates (duplicated by education and farm-urban variables), and (*c*) the black population percentage (duplicated by education and farm-urban variables).

Change in the black percentage is the best—and most important—example of hidden significance. Although a black variable is in five Percentage Change model equations, most equations (including the longest) lack a black variable. Yet no one who understands per-capita income change can fail to grasp the leading roles that (1) the high black percentages in former slave states had in causing the *low* 1950 per-capita incomes in those states—the lowest in 1950 per-capita income—and (2) South-to-North shifts in black population had in causing *change* in state per-capita income. Southern per-capita income change dominated the national per-capita income change picture. And southern change was in turn dominated by developments relating to blacks: the collapse of the black-oriented sharecropper system, improved education for southern blacks, and reduced employment discrimination against southern blacks. But in general, these developments are better represented in regression models by variables describing the specific developments (decline in farm and tenant farm population, increase in urban and metropolitan-area population, improved southern education, increased employment/population ratio). Even though the developments mentioned center on blacks and emphasize the South, they also have nonblack and non-South facets; general variables (covering blacks *and* whites) tend to have more explanatory power than black-specific ones. And when the general variables (e.g., median years of schooling) enter the equation, the black variables gradually lose significance.

The "alternate" equations include certain variables that usually, if not invariably, are not in the Best-*R* equations. These variables describe changes in the black population percentage, hourly wage rates, manufacturing, and the population percentage of widowed or divorced women. In the alternate equations, certain factors that overlap the ones being highlighted are not fully controlled. For example, the alternate eight-variable equation for the Percentage Change Model features POP%BLACK. This equation has only one education variable, whereas the best and second-best eight-variable equations have two (including MD-YRS-SCHOOL). Similarly, a six-variable equation featuring a wage change variable omits the college graduate variable, which tends to duplicate the hourly-wage-rate variable, and the equation has only one farm variable instead of two.

The alternate equations provide evidence that certain factors do affect per-capita income even though those factors are missing from the longest Best-*R* equations. One could argue, of course, that insignificance in the face of maximum controls (the longest equation) invalidates the hypotheses for these factors. But such an argument is not persuasive. Statistical findings must be interpreted in the light of what is known about the real-world situation. Given the knowledge that influences do overlap and that one influence can to some extent represent another, it would be reckless to ignore all statistical evidence other than that of the Best-*R* equations. The alternate equations impart balance to the analysis.

Slave-State and Consolidated Equations

The last three equations reported for each of the two basic models are (1) the highest *R* equation, (2) an augmented version of that equation that adds SLAVE (slave state dummy), and (3) a consolidated version that combines all variables representing the same factor (e.g., farm-urban mix or education) into one compound variable. The consolidated version omits SLAVE.

We use the augmented (SLAVE) version to eliminate ambiguity. The Best-*R* equations for the Percentage Change and State/Nation models contain interaction variables that use either SLAVE or NONSLAVE as multipliers; the multipliers respectively "dummy out" the former nonslave or slave states so that the resulting interaction variable describes variation only within a subset of states. The question arises: is the resulting interaction variable acting partly as a regional proxy—a proxy for SLAVE or NONSLAVE? This question can be answered by holding either SLAVE or NONSLAVE constant so that any *residual* tendency for the former slave states to have high or low growth in per-capita income is controlled. The "Plus SLAVE" equation answers the question by adding SLAVE to each of the Best-*R* equations. To anticipate, it turns out that SLAVE has no appreciable effect on the *t* values of the other variables. (The effects of slavery's legacy are adequately described by other variables, such as the agriculture and education variables.)

The Best-*R* equations have a second problem: multicollinearity, or duplication (intercorrelation) among variables. For example, the Best-*R* equation for the Percentage Change model has three farm-urban variables, two education variables, and two employment variables. The variables within each subset duplicate each other. The result is invalid—understated—*t* values for these seven variables. To get accurate significance readings on the overlapping variables, all variables representing a given factor must be combined into one variable, called a compound variable. The seven variables are reduced to three (one each for farm-urban mix, education, and employment). The resulting consolidated equation has six variables instead of ten. (Three of the original variables remain unchanged—unconsolidated.)

The procedure for computing the compound variables is that developed by Wheat:[1]

$$a + bW - cX + dY_1 - eY_2 - fY_3 + gZ_1 + hZ_2 =$$
$$a + bW - cX + 1.0(dY_1 - eY_2 - fY_3) + 1.0(gZ_1 + hZ_2)$$

where a is the regression constant; b–h are the regression coefficients from the original equation; W and X are nonduplicative variables (i.e., not being consolidated); Y_1, Y_2, and Y_3 are duplicative variables representing the Y factor; and Z_1 and Z_2 are duplicative variables representing the Z factor. To get the compound variables (enclosed in parentheses), you simply (1) multiply—weight—each of the variables being consolidated by its regression coefficient from the original equation and then (2) add the weighted variables for a given factor (Y or Z) together. As the algebraic transformation shows, the compound variable

always has a coefficient of 1.0. The variables not being consolidated (W and X) remain in the revised equation and retain their original coefficients (*b* and *c*). The constant (*a*) also remains the same. And consolidation has no effect on *R* or on the regression predictions. The only purpose of consolidation is to produce accurate significance tests (*t* values) for the factors whose variables are being combined. In other words, consolidation eliminates the effect on *t* of multicollinearity among two or more variables representing the same factor. Roughly speaking (but not in precise additive fashion), the compound Y variable's *t* adds together the *t*'s of Y's three constituent variables.[2]

Curvilinear Regression

We tested variables for curvilinear regression by comparing *t* values of the basic (linear) variable, its square, and its square root (the variable raised to the 0.5 power). If the basic variable was in a preliminary "best" equation but its square or square root was stronger when the equation's other variables were controlled, the stronger version was substituted in the final equation. Consequently, some equations have COLLEGE%^{2}, (ACRES/FARM)^2, and POP%WID-DIVOR$^{.5}$ in them. We used the linear versions of two of these variables—COLLEGE% and POP%WID-DIVOR—in control contexts (equations) where the tests for curvilinear regression revealed none.

Significance Levels

All equations are limited to variables that are significant at the 5 percent level or better. One-tail tests of significance are used, because we had theoretical expectations about the signs of the coefficients. For the shortest equations—two variables, 45 degrees of freedom—a *t* value of 1.68 is significant at the 5 percent level; a *t* of 2.42 is significant at the 1 percent level. For the longest equations—ten variables—the 5 percent and 1 percent significance levels are 1.69 and 2.43.

THE PERCENTAGE CHANGE MODEL

We begin with the stronger of the two basic models, the Percentage Change model. The dependent variable, once again, is the 1950–87 percentage increase in state per-capita income; both 1950 and 1987 use 1987 dollars. Five tables present the Best-*R* and alternate equations. The equations get progressively longer; each table after the first has longer equations than the preceding table. Our goal is to determine which factors influenced income change and what the relative importance of each factor is. The shortest equations, being the most restrictive, admit only the strongest factors. The longer a Best-*R* equation has to be before a certain factor is represented, the weaker that factor tends to be. Factors whose variables appear only in the longest equations (eight, nine, or ten

variables) tend to be the weakest. We also consider how variables rank by t in particular equations, how many variables a factor has representing it in each equation, and how many equations have the factor represented. And we consider the simple correlations (r's) from chapter 5.

Two-Variable and Three-Variable Equations

Table 6.1 shows the best and second-best two-variable (R_2) and three-variable (R_3) equations. These equations are the most restrictive. We again remind the reader that the independent variables describe 1950–87 *changes*, not base-year levels. The best R_2 equation uses POP%FARM (farm–urban mix) and COLLEGE%^{2} (education); the latter has a slightly higher t. Farm–urban mix and education variables also produced the two highest simple r's in the chapter 5 analysis. The two strongest factors by R_2 test and the two strongest by simple r test are thus the same. But whereas farm–urban mix won the simple r contest, the R_2 test puts education slightly ahead.

The second-best R_2 equation illustrates why researchers should not put too much stock in a single equation. In the second-best equation, education disappears from view; it is replaced by employment. The two variables are now SLxPOP%URBAN and POP%EMPLOYED. Why did education weaken? The answer lies in the switch from POP%FARM to SLxPOP%URBAN as the farm–urban variable. SLxPOP%URBAN has a stronger intercorrelation (+.46) with COLLEGE%^{2} than POP%FARM (+.06) has. Partial duplication of education by SLxPOP%URBAN weakens the education variables and allows the new factor, employment, to come in.

But education comes roaring back in the best R_3 equation. This equation is simply the best R_2 equation with MD-YRS-SCHOOL added. The three variables are POP%FARM, COLLEGE%^{2}, and MD-YRS-SCHOOL. With two variables in the equation, education is clearly the strongest factor. True, POP%FARM has the highest t, but that is because of collinearity (duplication) between the two education variables. Since COLLEGE%^{2} alone outperformed POP%FARM in the best R_2 equation, it is obvious that the *combination* of COLLEGE%^{2} and MD-YRS-SCHOOL has appreciably more explanatory power than POP%FARM in the R_3 equation.

The second-best R_3 equation adds COLLEGE%^{2} to the second-best R^2 equation. This addition results in an equation that embraces all three of the factors found in at least one of the preceding equations. The three factors, again, are farm–urban mix (SLxPOP%URBAN), education (COLLEGE%^{2}), and employment (POP%EMPLOYED). Education ranks third by t test; farm–urban mix ranks first. A consensus is emerging on what the three strongest factors are. But their relative strength remains uncertain.

Incidentally, SLxPOP%METRO (simple r = +.61) can replace SLxPOP%URBAN (+.69) in any equation; no appreciable reduction in R will result. The two variables have an intercorrelation of +.92.

Table 6.1
Percentage Change Model: Best and Second-Best Two-Variable (R_2) and Three-Variable (R_3) Equations

	Best R_2		Second-Best R_2		Best R_3		Second-Best R_3	
	Parameter	t	Parameter	t	Parameter	t	Parameter	t
Constant	-0.889	-3.32	0.915	16.10	-0.858	-3.60	0.618	6.67
POP%FARM	-1.883	-6.32			-1.679	-6.19		
SLxPOP%URBAN			1.314	7.42			1.011	5.78
COLLEGE%^{2}	0.179	6.71			0.131	4.81	0.102	3.79
MD-YRS-SCHOOL					0.880	3.58		
POP%EMPLOYED			1.386	4.17			1.440	4.93
R	.801		.789		.850		.846	
Adjusted R^2	.625		.605		.703		.696	

Four-Variable and Five-Variable Equations

Table 6.2 shows the best R_4 equation, an alternate R_4 equation featuring MFG/POP, the best R_5 equation, and the second-best R_5 equation. The best R_4 equation further extends the best R_2-R_3 pair of equations by adding UNEMPLOY%1987 (1987 unemployment rate). UNEMPLOY%1987 was intended to register the effects of short-run economic developments of the 1980s, but in practice it largely duplicates and replaces POP%EMPLOYED and UNEMPLOYED% (1950–87 change variables). Hence UNEMPLOY%1987 will be reinterpreted as an employment variable. In this equation, employment is the weakest of the three factors. The strongest factor, represented by two variables, is education.

The alternate R_4 equation, labeled MFG R_4, uses POP%FARM, COLLEGE%^{2}, TRANSCOSTDROP, and MFG/POP. The point of showing this equation is to illuminate the significance of the last two variables, particularly MFG/POP (1950–87 change in manufacturing employment per capita). MFG/POP has latent significance but tends to be duplicated by stronger variables. In particular, MFG/POP has a striking +.81 intercorrelation with POP%URBAN: controlling urban growth tends to control manufacturing growth. A further problem with MFG/POP is that manufacturing growth's effect in the West was negative. In the West, new manufacturing reduced the cost of living by eliminating long-distance shipping costs from the prices of manufactured goods that formerly were imported from the East. A lower price-wage structure with lower wages—hence lower income—resulted. The negative western effect tends to neutralize the positive effect of manufacturing growth elsewhere. The MFG R_4

Table 6.2
Percentage Change Model: Best and Alternate Four-Variable (R_4) and Five-Variable (R_5) Equations

	Best R_4		MFG R_4		Best R_5		Second-Best R_5	
	Parameter	t	Parameter	t	Parameter	t	Parameter	t
Constant	-0.629	-2.98	-0.844	-3.76	-0.461	-2.21	-0.523	-2.57
POP%FARM	-1.737	-7.46	-2.153	-8.36	-1.617	-7.25	-1.646	-7.39
POP%METRO							0.080	2.51
COLLEGE%[2]	0.123	5.25	0.123	4.84	0.114	5.14	0.107	4.69
MD-YRS-SCHOOL	1.062	4.94			1.040	5.14	0.957	4.62
UNEMPLOY%1987	-0.050	-4.13			-0.056	-4.82	-0.050	-4.41
TRANSCOSTDROP			-1.998	-4.56				
MFG/POP			0.230	3.23				
SLxPOP%BLACK					-0.436	-2.62		
R	.895		.874		.910		.909	
Adjusted R^2	.783		.742		.809		.806	

equation suggests that manufacturing growth really does have a net positive effect on income but that this effect is generally hidden in the Best-R equations. One Best-R equation, the best R_8 equation, does include MFG/POP: the present evidence does not stand alone.

The other new variable in the MFG R_4 equation is TRANSCOSTDROP, the western transport cost decline variable. High values of TRANSCOSTDROP signify big declines in the cost of shipping Manufacturing Belt goods to external states. We theorize that these declines reduced the cost of living in the West and thereby led to relatively lower wages and income. It is no coincidence that TRANSCOSTDROP and MFG/POP are in the same equation. Because manufacturing growth had a negative effect on income in the West, that negative western effect must be controlled for the positive relationship between MFG/POP and income in other regions to be seen. TRANSCOSTDROP indirectly controls MFG/POP's negative relationship for the West, because TRANSCOSTDROP incorporates the effect of increased manufacturing self-sufficiency in the West (one source of transport cost declines). With MFG/POP's negative relationship for the West controlled, the positive relationship found elsewhere can be seen. When we get to the best R_8 equation, where MFG/POP is again included, we shall see that TRANSCOSTDROP is again controlled. Meanwhile, the presence of TRANSCOSTDROP in the alternate R_4 equation gives us our first—but far from last—evidence that the transportation factor is significant.

Beyond manufacturing and transportation, the alternate R_4 equation adds to the

mounting evidence that farm-urban mix and education are the two strongest factors. In this equation the farm-urban variable, POP%FARM, is far ahead of the education variable, COLLEGE%^{2}. There is no employment variable.

The best R_5 equation is another expansion of the combination of POP%FARM and COLLEGE%^{2} that began with the best R_2 equation and continued through the best R_3 and R_4 equations. That is, the R_5 equation starts with the best R_4 equation and adds one more variable. The new variable is SLxPOP%BLACK. Like SLxPOP%URBAN, SLxPOP%BLACK "dummies out" the nonsouthern states by setting their values at zero; the variable thus focuses on declines in black population percentages in the former slave (SL) states. SLxPOP%BLACK's inclusion marks the first time a best or second-best equation has admitted a variable not representing one of the three strongest factors—farm-urban mix, education, and employment. The new (fourth) factor is racial mix. SLxPOP%BLACK's negative sign says that the southern states with the biggest declines in their black percentages tended to have the biggest increases income. That finding supports powerful theoretical considerations in affirming a racial mix influence. As already noted, POP%BLACK's intercorrelations with farm-urban and education variables—especially its -.85 r with SLxPOP%URBAN—keep it out of most equations.

The second-best R_5 equation is almost the same as the best; it merely substitutes POP%METRO for SLxPOP%BLACK. This substitution gives the equation two farm-urban variables to go with the two education variables and the one employment variable. We did not prepare consolidated versions of the shorter equations, but the equation's t values make it plain that the two strongest factors—farm-urban mix and education—are about equal in this equation. Employment ranks third. We should add that POP%METRO's presence here should dispel any doubts about whether the metropolitan facet of farm-urban mix bears on income growth. POP%METRO's significance jibes squarely with the well-established fact that metropolitan residents tend to have higher incomes.

Six-Variable Equations

Table 6.3 has four 6-variable equations: the best and second-best R_6 equations, an alternate equation labeled "WAGE R_6" that features WAGE-RATE, and another alternate equation that features POP%BLACK and TRANSCOSTDROP. The best R_6 equation starts with the best R_4 equation and adds two more farm-urban variables: NSxPOP%TENANT and SLxPOP%URBAN. A remarkable thing about this equation is that, despite being moderately long, it allows representation of just the three leading factors: farm-urban mix, education, and employment. Almost as remarkable is the fact that three of the six variables—POP%FARM, NSxPOP%TENANT, and SLxPOP%URBAN—are farm-urban variables. Whereas the best R_2, R_3, R_4, and R_5 equations ranked education ahead of farm-urban mix, this equation definitely puts farm-urban mix in the lead. Farm-urban mix has three variables to education's two; it has the best t (POP%FARM, t = -6.63); and its second-best t (4.10) is higher than education's second-best (3.87). Employment's failure to contribute more than one variable

Table 6.3
Percentage Change Model: Best, Second-Best, and Alternate Six-Variable (R_6) Equations

	Best R_6		Second-Best R_6		WAGE R_6		TRANSCOST DROP & POP%BLACK R_6	
	Parameter	t	Parameter	t	Parameter	t	Parameter	t
Constant	-0.574	-2.35	-0.708	-3.07	-0.155	-0.84	0.571	4.47
POP%FARM	-1.512	-6.63	-1.483	-6.65				
NSxPOP%FARM					-0.587	-4.49	-0.323	-2.54
POP%TENANTFARM			-0.574	-2.16				
NSxPOP%TENANT	-0.412	-3.51						
SLxPOP%URBAN	1.101	4.10			1.900	7.15	1.060	2.79
COLLEGE%^{2}	0.102	4.82	0.118	5.52			0.083	3.58
MD-YRS-SCHOOL	0.865	3.87	0.756	3.23	1.153	4.15		
UNEMPLOY%1987	-0.062	-5.75	-0.060	-5.33				
POP%EMPLOYED							1.337	5.24
POP%LABOR					1.412	3.53		
UNEMPLOYED%					-0.164	-3.94		
POP%BLACK							-0.215	-2.32
SLxPOP%BLACK			-0.433	-2.71				
WAGE-RATE					0.637	2.67		
TRANSCOSTDROP							-13.958	-3.67
R	.927		.920		.905		.901	
Adjusted R^2	.840		.824		.792		.785	

strengthens the evidence that employment ranks behind farm–urban mix and education. This finding is softened by the strengthened evidence that employment nevertheless outranks all the remaining factors (not in the equation).

The presence of NSxPOP%TENANT (NONSLAVE x POP%TENANT-FARM) in the best R_6 equation is intriguing. Because of the extreme poverty of *southern* tenant farmers (this group includes sharecroppers), we thought that the southern version of the variable (SLxPOP%TENANT) was more likely to enter some equations. But the nonsouthern version (NSxPOP%TENANT) does have its own logic. Nonsouthern tenant farmers, though not in the same league as southern ones, have much lower incomes than nonsouthern farmers living on owner-

occupied farms. Once the income gains caused by decline in the general farm population have been controlled by POP%FARM (also in the equation), additional income gains can be explained by decline in the *tenant* farm population. Why doesn't the tenant farmer effect register for the South too? The probable reason is that SLxPOP%URBAN is in the equation. Its intercorrelation of -.90 with SLxPOP%TENANT makes it a good proxy for the latter: tenant-sharecropper decrease and urban increase in the South were two aspects of the same farm-to-urban population shift phenomenon.

Table 6.3's second-best R_6 equation provides further evidence of the effect of racial mix changes. The only difference between the best and second-best R_6 equations is that the second-best one (*a*) replaces SLxPOP%URBAN and NSxPOP%TENANT with one variable, POP%TENANTFARM, and (*b*) adds SLxPOP%BLACK. Because of SLxPOP%URBAN's -.90 intercorrelation with SLxPOP%TENANT, the first change—change (*a*)—is almost the equivalent of replacing SLxPOP%TENANT and NSxPOP%TENANT with the combined variable, POP%TENANTFARM. That consolidation makes room for SLxPOP%BLACK. R declines only from .927 to .920. The effectiveness of SLxPOP%BLACK as a substitute variable reiterates that black out-migration from the South stimulated southern income growth. The absence of black population variables from most equations reflects duplication of racial mix by farm-urban and education variables, not a lack of importance for race.

The equation labeled WAGE R_6 is included to show that WAGE-RATE (1950–87 change in manufacturing's hourly wage rate) *is* significant. This evidence supports our earlier suggestions that the wage factor is significant and that its absence from other equations results from duplication by other variables. The education and farm-urban variables are the chief duplicators; all three factors—education, farm-urban mix, and hourly wages—had their biggest changes in the South. Note that the WAGE R_6 equation has only one education variable, compared to the two found in both the best and second-best R_6 equations. Likewise, the WAGE R_6 equation has no variable describing farm or tenant farm population decline in the South; the only farm variable, NSxPOP%FARM, has explanatory power only for the nonsouthern states.

Two other features of the WAGE R_6 equation inspire comment. First, the employment factor is strong enough to bump education down to third rank in some equations. In this equation, employment has two variables to education's one. The near equality of the t's for MD-YRS-SCHOOL (4.15) and UNEMPLOYED% (-3.94) leaves no doubt that the second employment variable, POP%LABOR (3.53), carries enough weight to put employment ahead of education. This evidence counterbalances other evidence suggesting that education is the strongest factor. Second, the slave (SL) and nonslave (NS) states seem to emphasize different sides of the farm-urban dichotomy. The southern (slave) states choose an urbanization variable, SLxPOP%URBAN, to represent them, but the nonslave (nonsouthern) states choose a farm variable, NSxPOP%FARM. Part of the explanation seems to be that SLxPOP%URBAN partly represents POP%BLACK, with which it has a -.85 intercorrelation. The highest POP%BLACK values (positive) and the lowest SLxPOP%URBAN values (zero) are in the nonsouthern states; the lowest POP%BLACK values (negative) and the highest

SLxPOP%URBAN values (positive) are in the southern states. The farm side of the farm–urban dichotomy may be emphasized outside the South because of the major contribution farm population decline made to income growth in the Plains.

Table 6.3's last equation, TRANSCOSTDROP & POP%BLACK R_6, provides further evidence that changes in racial mix (POP%BLACK) caused income change despite the absence of a black population percentage variable from most equations. Secondarily, the equation offers additional evidence of the significance of TRANSCOSTDROP, which is absent from the shorter Best-R equations. Although POP%BLACK is the equation's least significant variable (t = -2.32), that weakness results because SLxPOP%URBAN is controlled. We just saw that POP%BLACK and SLxPOP%URBAN have a -.85 intercorrelation. POP%BLACK's ability to enter the equation in the face of that much duplication is really a demonstration of strength. One reason POP%BLACK can enter this equation but not most of the others is that this equation has only one education variable controlled. MD-YRS-SCHOOL, with which POP%BLACK has a -.60 intercorrelation, is missing. (POP%BLACK's intercorrelation with COLLEGE%^{2} is only - .39.) To anticipate, MD-YRS-SCHOOL is also absent from the BLACK R_8 equation, which again features POP%BLACK.

Little needs to be said about TRANSCOSTDROP. Its absence from most previous equations results from the greater importance of education, farm–urban mix, and employment. Those factors are so strong that they dominate the shorter equations, squeezing the less important factors out. All TRANSCOSTDROP needs to display its significance is a long enough equation. It will be in every equation from here on.

Eight-Variable and Nine-Variable Equations

Table 6.4 presents the best and second-best eight-variable equations, an alternate eight-variable equation (BLACK R_8) that again features POP%BLACK, and the best nine-variable equation. In the best R_8 equation, farm–urban mix and education are in their customary positions as the two strongest factors; employment again ranks third. But the revealing aspect of this equation is the presence of four variables representing lesser factors: SLxPOP%BLACK (racial mix), MFG/POP (industry mix), TRANSCOSTDROP (cost of shipping eastern goods to the West), and POP%WID-DIVOR (widowed-divorced female population percentage). SLxPOP%BLACK previously appeared in the best R_5 and second-best R_6 equations. But with those two exceptions, no previous best or second-best equation has had a variable for any factor other than the three strongest. (TRANSCOSTDROP was in two alternate equations and MFG/POP in one, but these equations were not the best or second-best for the given number of variables.)

The four lesser factors, slighted until now, deserve attention. SLxPOP%BLACK, negative, adds to the evidence that black population percentage declines in the South (former slave—SL—states) stimulated southern income growth. This variable has now been in three best or second-best equations. MFG/POP, change in manufacturing employment per capita, is

Table 6.4
Percentage Change Model: Best and Alternate Eight-Variable (R_8) and Best Nine-Variable (R_9) Equations

	Best R_8		Second-Best R_8		BLACK R_8		Best R_9	
	Parameter	t	Parameter	t	Parameter	t	Parameter	t
Constant	-0.585	-2.39	-0.838	-2.82	-0.062	-0.22	-0.752	-2.67
POP%FARM	-1.801	-7.86	-0.901	-3.13			-0.917	-3.38
NSxPOP%FARM					-0.365	-3.00		
POP%TENANTFARM					-0.578	-2.03		
NSxTENANT%xAC			-0.273	-3.90			-0.311	-4.59
SLxPOP%URBAN			1.197	4.61	0.898	2.51	1.195	4.89
COLLEGE%	0.323	4.57	0.373	5.07			0.384	5.54
COLLEGE%^{2}					0.087	3.87		
MD-YRS-SCHOOL	0.846	3.55	0.565	2.08			0.805	2.94
UNEMPLOY%1987	-0.042	-3.48						
POP%EMPLOYED			1.132	3.80	1.298	5.46	0.974	3.38
POP%BLACK					-0.221	-2.53		
SLxPOP%BLACK	-0.419	-2.60						
MFG/POP	0.172	2.73						
TRANSCOSTDROP	-10.59	-2.35	-11.175	-2.74	-14.825	-4.06	-11.631	-3.03
POP%WID-DIVOR	-0.504	-2.34					-0.530	-2.46
OLD-INMIGRATN			10.359	2.05			13.035	2.67
AMENITIES					0.089	1.85		
R	.930		.928		.920		.938	
Adjusted R^2	.838		.833		.815		.852	

positive: income growth tended to be high in states with high manufacturing growth. But there is an important qualification. As in the MFG R_4 equation, MFG/POP's significance depends on TRANSCOSTDROP's being controlled. Manufacturing growth in the West had a negative effect on income, because it reduced the need to import eastern manufactured goods with their high long-distance shipping costs. This effect was part of the West's transport cost reduction effect, which lowered the West's cost of living and thereby lowered wages and income. With the transport cost decline controlled, MFG/POP becomes significant.

TRANSCOSTDROP and POP%WID-DIVOR are the two weakest variables in the equation; their *t* values (-2.35 and -2.34) are barely short of the 1 percent significance level, 2.43. But subsequent equations will give them significance well above the 1 percent level. For that matter, TRANSCOSTDROP has already displayed *t*'s of -4.56 and -3.67 in alternate R_4 and R_6 equations. TRANSCOSTDROP's negative sign declares that states with high costs of shipping goods to them from the Manufacturing Belt in 1950—costs that declined sharply between 1950 and 1987—tended to have low increases (relative declines) in income. The negative sign for POP%WID-DIVOR signifies that high increases in the percentage of widowed and divorced women were associated with low increases in income.

The second-best R_8 equation gives a one-two-three ranking for farm-urban mix (three variables), education (two variables), and employment (one variable). For these three factors, the most interesting innovation in the equation is the emergence of a farm acreage (average farm size) influence, a special facet of farm-urban mix. NSxTENANT%xAC is a double interaction variable: NONSLAVE x POP%TENANTFARM x ACRES/FARM. Since POP%TENANTFARM always has negative values and ACRES/FARM always has positive values, POP%TENANTFARM x ACRES/FARM is always negative. The mathematical effect of multiplying them is to increase the relative magnitude of the tenant farm population decrease in the states that had the biggest increases in average farm acreage. The substantive effect is to say that the income gain resulting from tenant decline is magnified in whatever degree the remaining farmers have larger farms, hence larger incomes. The connection between the two developments—the basis for viewing them as interacting—is that the vacated tenant farms got consolidated with the remaining farms. The overall variable, NSxTENANT%xAC, highlights the tenant-acres interaction in the nonsouthern states; the southern states (NS = 0) get dummied out—valued at zero.

Only two variables not belonging to the three leading factors are in the second-best R_8 equation. The variables are TRANSCOSTDROP and OLD-INMIGRATN. TRANSCOSTDROP's *t* (-2.74) is now safely past the 1 percent level (2.43). OLD-INMIGRATN (1975–80 in-migration rate, persons 60 and older) is significant at only about the 3 percent level. But OLD-INMIGRATN's *t* will improve when more variables are controlled; the more important fact for this equation is that *t* is significant at all. This is the first equation in which OLD-INMIGRATN appears. We thus have the first evidence that income growth was stimulated in certain states by an influx of affluent retirees.

Table 6.4's BLACK R_8 equation is included to display two normally absent variables—POP%BLACK and AMENITIES. POP%BLACK is normally absent because it is duplicated by the education variables, which are stronger because they describe the entire population. But in this equation MD-YRS-SCHOOL, one of the two usual education variables, is not present—not controlled. MD-YRS-SCHOOL's -.60 intercorrelation with POP%BLACK militates against efforts of POP%BLACK to enter equations containing MD-YRS-SCHOOL. POP%BLACK's presence in this equation counters the suggestion of SLxPOP%BLACK in other equations that black population change might have affected income only in the South. (When we get to the State/Nation model, we

will present two equations—one is the best overall equation—where NSxPOP%BLACK is included.)

AMENITIES is the other new variable. It is a dummy valued at one in 19 states where seacoasts, lakes, and mountains have created large retiree populations in amenity-rich counties. The theory behind this variable is that amenities directly and indirectly stimulate income. Directly, they attract affluent retirees. Indirectly, they attract local-market-oriented manufacturing, which tends to grow in response to population growth. (Local-market-oriented firms sell most of their output in the city or metropolitan area where they are located; examples are newspapers, bakeries, and bottling plants.) The link between amenities and retiree migration is seen in the fact that t for OLD-INMIGRATN declines by 0.52 when AMENITIES enters the equation. Similarly, in the second-best R_8 equation, the partial r for AMENITIES (not in the equation) declines by 0.52 in its t when OLD-INMIGRATN enters. This relationship and the data used to construct AMENITIES lead us to interpret it as another retiree migration variable. AMENITIES adds to the evidence from OLD-INMIGRATN that retiree migration stimulates income.

The BLACK R_8 equation also illuminates three additional points. First, employment (POP%EMPLOYED) is strong enough in this equation, as in three previous ones, to rank second in importance. This finding strengthens employment's grip on third place (its usual rank) in the overall consensus of evidence. Second, education ranks only fourth in this equation; even TRANSCOSTDROP is stronger here than COLLEGE%^{2}. This evidence counterbalances evidence from other equations that education is a stronger influence than farm–urban mix. Third, transportation (TRANSCOSTDROP) is a consistently significant and sometimes strong factor (third-ranked here) in the longer equations.

The best R_9 equation is of interest primarily for four of its variables: NSxTENANT%xAC, TRANSCOSTDROP, POP%WID-DIVOR, and OLD-INMIGRATN. All of these except TRANSCOSTDROP have appeared in a previous equation only once. And two of them (POP%WID-DIVOR and OLD-INMIGRATN) fell a bit short of the 1 percent level in their previous appearances. The new evidence confirms the significance of all four variables and the influence of the factors they represent. (In the case of NSxTENANT%xAC, what needs confirmation is really a subfactor: farm size, or the AC—acreage—part of the variable.) Among the three lesser factors, transportation ranks highest. This is not surprising: TRANSCOSTDROP is in five previous equations, not just one.

Best Overall Equation

Table 6.5 has three versions of the ultimate equation: the best ten-variable(R_{10}) equation. The variables in this equation yield the highest R (.943) attainable by using significant (at the 5 percent level), right-sign variables. Actually, all variables except one (UNEMPLOYED%) are significant at or essentially at the 1 percent level. POP%EMPLOYED (t = 2.36) does fall a hairsbreadth short of the 1 percent level (2.43), but the minute deficiency results from duplication. The

Table 6.5
Percentage Change Model: Best Ten-Variable (R_{10}) Equation: Basic, SLAVE, and Consolidated Versions

	Basic Best R_{10}		Best R_{10} plus SLAVE		Consolidated Best R_{10}	
	Parameter	t	Parameter	t	Parameter	t
Constant	-0.580	-1.99	-0.606	-1.85	-0.580	-4.03
POP%FARM	-0.773	-2.79	-0.773	-2.76		
NSxTENANT%xAC	-0.332	-4.96	-0.350	-2.93		
SLxPOP%URBAN	1.368	5.30	1.352	4.89		
* FARM-URBAN					1.000	8.72
COLLEGE%	0.370	5.44	0.374	5.18		
MD-YRS-SCHOOL	0.877	3.25	0.861	3.00		
* EDUCATION					1.000	9.39
POP%EMPLOYED	0.737	2.36	0.735	2.33		
UNEMPLOYED%	-0.083	-1.75	-0.086	-1.70		
* EMPLOYMENT					1.000	5.33
TRANSCOSTDROP	-11.63	-3.11	-11.68	-3.07	-11.63	-3.55
POP%WID-DIVOR	-0.564	-2.67	-0.556	-2.56	-0.564	-3.07
OLD-INMIGRATN	12.500	2.62	12.439	2.57	12.500	2.87
SLAVE			0.032	0.18		
R		.943		.943		.943
Adjusted R^2		.859		.856		.873

* Compound variable (explained in text). FARM-URBAN = -0.773 POP%FARM - 0.332 NSxTENANT%xAC + 1.368 SLxPOP%URBAN; EDUCATION = 0.370 COLLEGE% + 0.877 MD-YRS-SCHOOL; EMPLOYMENT = 0.737 POP%EMPLOYED - 0.083 UNEMPLOYED%.

R_{10} equation is simply the best R_9 equation with UNEMPLOYED% added. In the R_9 equation POP%EMPLOYED has a formidable t of 3.38, but this t declines to 2.36 when the second employment variable is added; POP%EMPLOYED's significance cannot be doubted.

An adjusted R^2 of .859 says that the equation explains 86 percent of the variation in 1950–87 state per-capita income growth. The equation would explain even more of the income variation were it not for time period mismatches between the dependent variable (1950–87) and seven of the ten independent variables (1950–80 or 1975–80). All five of the most important variables—the three farm-urban mix variables and the two education variables—use 1950–80.

And OLD-INMIGRATN may be seriously weakened by having to use 1975–80 migration as a proxy for 1950–87 migration.

The equation's three versions are (1) the basic ten-variable equation, (2) an augmented version that adds an eleventh variable, SLAVE, and (3) the consolidated version, which combines all variables representing a given factor into one compound variable. The consolidated equation has six variables: three compound variables and three ordinary (noncompound) variables. The augmented version with the slave state dummy serves to verify that the two interaction variables—SLxPOP%URBAN and NSxTENANT%xAC—are not essentially proxies for their dummy multipliers (SLAVE and NONSLAVE, or SL and NS). The consolidated version allows factors to be ranked by *t* value; it eliminates the effects of duplication within subsets of two or three variables that describe the same factor (farm–urban mix, education, or employment).

The Basic Equation

Look first at the basic equation. The first thing to note is that the usual three strongest factors—farm–urban mix, education, and employment—are still the three strongest. There are (*a*) three farm–urban variables, including the equation's second-strongest and third-strongest variables, (*b*) two education variables, including the equation's strongest variable, and (*c*) two employment variables. Employment is definitely the weakest (third-ranked) of the three factors: its strongest *t* (2.36) is considerably outweighed by the strongest of both farm–urban mix (5.30) and education (5.44), and its second *t* (-1.75) is likewise outweighed by farm–urban mix's (-4.96) and education's (3.25). Incidentally, the second-strongest farm–urban variable is NSxTENANT%xAC—the variable that interacts tenant farm population decrease with average farm size increase. This variable is now in three equations, and its *t* of -4.96 is highly significant. From this evidence we conclude that farm acreage (AC) increases did affect income, even though the effect shows up mainly in interaction variables. (A noninteraction farm size variable is significant in one State/Nation model equation.)

Among the three remaining variables, TRANSCOSTDROP is strongest; its *t* of -3.11 is well beyond the 1 percent significance level. This performance, following similar performances in six previous equations, strongly supports the transport cost decline theory of relative income decline for the West. The last two variables—POP%WID-DIVOR and OLD-INMIGRATN—are virtually tied; their respective *t*'s are -2.67 and 2.62. Both variables were in two previous equations, and both reach the 1 percent significance level. We can be reasonably confident, therefore, that their significance is genuine. This not to deny that the factors they represent seem to be the two weakest.

The "Plus SLAVE" Equation

Table 6.5's second equation is labeled "Best R_{10} Plus SLAVE." This 11-variable equation simply adds the slave-state dummy (essentially a South dummy) to the previous equation. The new equation's purpose is to determine whether the two interaction variables—SLxPOP%URBAN and NSxTENANT%xAC—really

represent their substantive parts: urbanization (POP%URBAN) and tenant farming (POP%TENANT FARM). The alternative possibility is that the interaction variables (or possibly just one) get most of their explanatory power from their dummy multipliers—SLAVE (SL) and NONSLAVE (NS). If the two interaction variables are essentially proxies for the dummies, SLxPOP%URBAN and NSxTENANT%xAC should lose their significance when SLAVE is controlled. (NONSLAVE does not have to be controlled, because it has a -1.00 intercorrelation with SLAVE: states valued at one and states valued at zero for SLAVE are respectively valued at zero and one for NONSLAVE.)

The SLAVE equation shows that controlling SLAVE has virtually no effect on SLxPOP%URBAN: SLxPOP%URBAN's *t* slips only from 5.30 (10-variable equation) to 4.89 (11-variable equation), remaining the second-highest *t*. NSxTENANT%xAC's *t* declines from -4.96 to -2.93, or roughly from 5 to 3. The new *t* is still well above the 1 percent level. This outcome means that SLxTENANT%xAC is not primarily a proxy for SLAVE but instead mainly describes the effect of tenant farm population declines (interacting with average acreage increases) in the nonsouthern states.

The Consolidated Equation

Table 6.5's third equation, Consolidated Best R_{10}, combines the basic equation's three farm-urban variables into one compound variable (FARM-URBAN), its two education variables into one (EDUCATION), and its two employment variables into one (EMPLOYMENT). The resulting equation has six variables—the three compound variables (FARM-URBAN, EDUCATION, and EMPLOYMENT) plus TRANSCOSTDROP, POP%WID-DIVOR, and OLD-INMIGRATN.

The purpose of consolidation is to deal with multicollinearity, that is, to eliminate the depressing effects on *t* of having two or three substantively overlapping variables measuring the same general factor. As explained earlier, consolidation pools the explanatory power of all variables measuring the same factor into one variable—and one *t*. The result is a reliable *t*, one that is not understated. Given just one variable per factor, we can rank the factors in importance by ranking the variables that represent them. The footnote to table 6.5 gives the formulas for the three compound variables; the general formula for compound variables is discussed in this chapter's first section.

The basic equation establishes beyond reasonable doubt that, where the R_{10} equation is concerned, farm-urban mix and education are the two strongest factors. Employment obviously ranks no better than third: the stronger of the two employment variables has a lower *t* (2.36) than the weakest of the three farm-urban variables (-2.79) and the weaker of the two education variables (3.25). The chief question to be answered by the consolidated equation thus is: Is farm-urban mix or education the stronger factor in the R_{10} equation? Another question is: Does employment still outrank transportation?

The first question's answer is that education outranks farm-urban mix in the basic equation. The comparative *t* values for the three compound variables (factor variables) are 9.39 for EDUCATION, 8.72 for FARM-URBAN, and 5.33 for

EMPLOYMENT. EDUCATION's primacy is mildly surprising, because the basic equation has three farm-urban variables to education's two. But when we note that the most significant variable in the basic equation is an education variable, COLLEGE%, the outcome is understandable.

The second question's answer is that EMPLOYMENT (5.33) easily outranks TRANSCOSTDROP (-3.55), which in turn outranks POP%WID-DIVOR (-3.07) and OLD-INMIGRATN (2.87). This result follows the drift of the many previous equations. Those equations suggest that employment has a firm grip on third place in the importance rankings.

Consensus of Percentage Change Model Evidence

Although the R_{10} equation entails the most control and is thus the most reliable, no one equation is definitive. And neither should the correlation evidence be ignored. So let us pull together the findings of the 17 equations that use INCOME/POP:% as the dependent variable and mix in the correlation evidence. Our goal is to rank by importance the factors that influenced income. Our analysis points to the following ranking: (1 and 2) education and farm-urban mix, tied, (3) employment, (4) racial mix, (5) transportation, (6) the widowed-divorced female percentage, (7) retiree migration, (8) industry mix, or manufacturing, and (9) hourly wage rates.

The Top Three Factors

Education, farm-urban mix, and employment are undoubtedly the three leading factors. Education and farm-urban mix are neck and neck in the race for first place. But education *seems* to come out slightly ahead. Education ranks first not only in the best overall (R_{10}) equation but in the four most restrictive Best-R equations: the best R_2, R_3, R_4, and R_5, equations. In addition, the simple r comparisons give an education variable, MD-YRS-SCHOOL, the highest r (+.62) of any noninteraction variable. (We discount the regional interaction variables, because their regional dummy components have strong proxy characteristics until other things are controlled.) The strongest farm-urban variable, POP%FARM (-.53), is outranked not only by education variables but by race and wage variables.

On the other hand, some equations give the edge to farm-urban mix or are too close to call. The best R_6 equation gives the edge to farm-urban mix. So do the second-best R_2, R_3, and R_5 equations; the MFG R equation; the WAGE R equation; and the BLACK R_6 and R_8 equations. In two of these eight equations, education falls to third rank, and in two others—the BLACK equations—it falls to fourth rank; whereas farm-urban mix never ranks below second.

Moreover, there is good evidence that the education variables represent race as well as education. MD-YRS-SCHOOL has a -.60 intercorrelation with POP%BLACK, and POP%BLACK's simple r of -.61 marks racial mix as a strong enough influence to give education a boost. (The only noninteraction r higher than POP%BLACK's -.61 is MD-YRS-SCHOOL's -.62.) Predictably,

education drops to fourth in the two equations that control POP%BLACK. And in one-variable-held-constant tests that control only POP%BLACK, the partial *r*'s for the best farm-urban and education variables are similar: -.39 for POP%FARM (down from -.53), +.40 for MD-YRS-SCHOOL (down from +.62), and +.45 for COLLEGE%^{2} (down from +.57).

Given the conflicting evidence, we think the most justifiable conclusion is that education and farm-urban mix are tied for first place—equal in importance. The consolidated R_{10} equation's *t*'s of 9.39 for EDUCATION and 8.72 for FARM-URBAN are so close that the difference fails to justify a firm conclusion that one influence is stronger than the other. Likewise, the *r*'s of 6.71 for COLLEGE%^{2} and -6.32 for POP%FARM in the most restrictive best-*R* equation, the R_2 equation, are too close to give education a clear-cut advantage. And we recognize that education has a stronger tendency to represent racial mix secondarily than farm-urban mix has: the strongest farm-urban variable by simple *r* test, POP%FARM, has an intercorrelation of only +.42 with POP%BLACK, compared to that of -.60 between MD-YRS-SCHOOL and POP%BLACK.

Employment (POP%EMPLOYED, POP%LABOR, UNEMPLOYED%, and UNEMPLOY%1987) is securely ensconced in third place. EMPLOYMENT's *t* of 5.33 (supported by *two* variables) in the consolidated R_{10} equation puts it far ahead of the other contenders for third place. Employment variables are the only variables other than farm-urban and education variables to be found in the best or second-best R_2, R_3, and R_4 equations, as well as the second-best R_5 and best R_6 equations. The best R_5 and second-best R_6 equations include a fourth factor—racial mix—but employment continues to rank third. It again ranks third in the R_8 and R_9 equations, which introduce still more factors. And in the BLACK R_6 and R_8 equations, POP%EMPLOYED has the highest *t* and overshadows the lone education variable (COLLEGE%^{2}), although farm-urban mix ranks first on the strength of multiple variables.

The Mid-Rank Factors: Racial Mix and Transportation

Racial mix (POP%BLACK and SLxPOP%BLACK) seems to come out slightly ahead in the competition for fourth place. Only two best or second-best equations through R_6 include a variable that doesn't represent farm-urban mix, education, or employment. Those two equations—best R_5 and second-best R_6—have just one such additional variable: SLxPOP%BLACK. SLxPOP%BLACK is also in the best R_8 equation, although there it ranks fifth (behind industry mix) when the *factors* are ranked. Another race variable, POP%BLACK, is in two alternate equations, where it ranks fifth (behind transportation). Heavy duplication of racial mix by education and farm-urban mix squeezes racial mix out of the other equations. But it bears repeating that the second-best simple *r* for a noninteraction variable belongs to POP%BLACK (+.61).

Our conclusion that racial mix ranks fourth also incorporates the plain facts of recent economic history. Sharecropping, a black-oriented institution, was the keystone of southern poverty. Sharecropping's collapse resulted in the migration of millions of blacks from the South to other regions. This migration (1) brought an interregional transfer of poverty, (2) gave the South an exceptionally large

farm-to-urban population shift, (3) helped raise average educational attainment in the South, and (4) restrained growth in per-capita employment outside the South. The farm–urban and education variables, and to some extent the employment variables, capture these developments, preventing POP%BLACK from receiving due recognition in the regression tests. This proxy representation of racial mix must be considered when we interpret the regression results. We think it inconceivable that changes in racial mix had less effect on income than the factor we rank next, transportation. And we think the comparative simple r^2s of .37 for POP%BLACK and .13 for TRANSCOSTDROP amply support our opinion.

Transportation has a firm grip on fifth rank. No best or second-best equation embraces more than the top four factors until we get to eight-variable equations. Both the best and second-best R_8 equations include TRANSCOSTDROP; transportation is the only factor other than the top three that appears in both equations. TRANSCOSTDROP is also in three alternate equations: MFG R_4, BLACK R_6, and BLACK R_8. In the MFG R_4 equation, it is the third-strongest variable. In the best R_9 and best R_{10} equations, the transportation factor ranks fourth—ahead of every other factor except the top three. TRANSCOSTDROP's simple r of -.36 is significant at the 1 percent level; the simple r's of POP%WID-DIVOR and OLD-INMIGRATN, representing the sixth-ranked and seventh-ranked factors, are close to zero.

The Four Least Important Factors

Next in importance are changes in the widowed-divorced female percentage of total population (POP%WID-DIVOR) and migration by affluent retirees (OLD-INMIGRATN and AMENITIES). POP%WID-DIVOR and OLD-INMIGRATN are the two least significant variables in the best R_9 and R_{10} equations. OLD-INMIGRATN is ahead in the R_9 equation, POP%WID-DIVOR in the R_{10} equation. Our judgment on which ranks highest therefore rests on the remaining two equations where these two variables appear: the best and second-best R_8 equations. Only POP%WID-DIVOR is in the best R_8 equation, where it is significant at about the 1 percent level. Only OLD-INMIGRATN is in the second-best R_8 equation, but it reaches only the 3 percent level. Because POP%WID-DIVOR is in the best equation and reaches the 1 percent level, we rank it higher—sixth. Retiree migration ranks seventh.

Eighth rank goes to industry mix, which really boils down to manufacturing. No industry variables other than manufacturing variables enter any equations. Manufacturing (MFG/POP) is in only one best or second-best equation: the best R_8 equation, where industry mix ranks fourth. MFG/POP is also in the MFG R_4 equation, but that equation must be discounted; its main purpose is to verify that manufacturing was a significant influence. Although MFG/POP has a simple r of only +.15 (insignificant), HIMID-MFG/POP reaches +.41—comfortably above the 1 percent level and also above TRANSCOSTDROP's -.36.

Ninth (and last) rank belongs to hourly wage rate increases (WAGE-RATE). WAGE-RATE is in only the subjectively chosen WAGE R_6 equation. We show this equation to demonstrate that WAGE-RATE does achieve significance in the face of fairly substantial controls. But WAGE-RATE also has an extremely im-

pressive simple *r* of +.59. This *r* is exceeded among the noninteraction *r*'s only by MD-YRS-SCHOOL's +.62 and POP%BLACK's -.61. The reason WAGE-RATE weakens and ultimately becomes insignificant is that, like POP%BLACK, it is cumulatively duplicated by other variables. The biggest wage gains were in the South. The South also had the biggest gains in urban population percentage and education; it had the biggest declines in farm, tenant farm, and black population. The biggest wage declines were in the West. The West also had the biggest per-capita employment increases and transportation cost (and cost of living) decreases. Variables representing the five other factors thus indirectly represent wages.

It is statistically possible, of course, that WAGE-RATE is never anything but a proxy for other variables that it overlaps. But that statistical possibility is so implausible that we can't take it seriously. To begin with, if WAGE-RATE (+.59) were nothing but a proxy, it could hardly have almost the same simple *r* as MD-YRS-SCHOOL (+.62) and POP%BLACK (-.61): WAGE-RATE has intercorrelations of only +.57 and -.42 with those variables—far short of what would be needed for them to carry WAGE-RATE to +.59. And how do we explain why WAGE-RATE's +.59 exceeds POP%FARM's -.53, if WAGE-RATE is nothing but a proxy? Theoretical considerations also argue that WAGE-RATE's significance is genuine. How could the impressive wage gains of the South and the large declines in relative wages in the West not have affected income? Rather than doubting whether the wage influence should be ranked at all, we doubt that it deserves to be ranked below the preceding three or four factors. Fifth or sixth might be a more accurate ranking, even if the evidence suggests otherwise.

THE STATE/NATION MODEL

In the State/Nation model, the dependent variable is INCOME/POP: S/N, the 1950–87 absolute change in the ratio of state to national per-capita income (i.e., in state per-capita income as a percentage of national per-capita income). The independent variables follow the same State/Nation (S/N) change formula. To avoid lengthening and complicating the labels for the independent variables, we use the same labels for both models—Percentage Change and State/Nation. But each model calculates the variables differently.

The State/Nation model's equations are remarkably similar to the Percentage Change model's. But there are differences. First, employment (POP%EMPLOYED, POP%LABOR, and UNEMPLOYED%) has more importance in the State/Nation model, particularly in the shorter equations. Second, the urban side of farm-urban mix is a little stronger, and SLxPOP%METRO replaces SLxPOP%URBAN as the dominant urbanization variable. Third, the farm side of farm-urban mix is a little weaker and puts more emphasis on *nonsouthern* farm population decline (NSxPOP%FARM, NSxPOP%TENANT) than found in the Percentage Change model. Fourth, increase in average acres per farm (ACRES/FARM), seen only in interaction variables in the Percentage Change model, is a separate variable in one State/

Nation equation—in addition to being part of an interaction variable in another equation. Fifth, NSxPOP%BLACK, which spotlights *nonsouthern* changes (increases) in the black population percentage, is now significant; in the Percentage Change model, SLxPOP%BLACK, which spotlights *southern* changes (decreases), was in three equations. Sixth, MFG/POP (change in manufacturing employment per capita), found in some Percentage Change equations, is not in any State/Nation equation. Seventh, UNEMPLOY%1987 is likewise missing from the State/Nation model. Eighth, OLD-INMIGRATN is also missing.

Because of the State/Nation model's generally similar results and somewhat weaker predictive power, we will discuss the State/Nation findings in less detail than we provided for the Percentage Change model.

Two-Variable and Three-Variable Equations

Table 6.6 shows the best and second-best two-variable and three-variable equations for the State/Nation model. The dominance of the three leading factors continues: the only factors represented in any of these first four equations—the most restrictive equations—are farm-urban mix, education, and employment. The Best R_2 equation has an education variable (MD-YRS-SCHOOL) and an employment variable (POP%EMPLOYED). As in the Percentage Change model's best R_2 equation, the education variable is more significant. But there the resemblance ends: the earlier model's farm-urban variable (POP%FARM) has been replaced by the employment variable (POP%EMPLOYED).

The neglect of farm-urban mix is only temporary, however. In the second-best R_2 equation, FARM%xACRES (POP%FARM x ACRES/FARM) not only replaces MD-YRS-SCHOOL but is the equation's strongest variable. The presence of the farm size (ACRES/FARM) facet of farm-urban mix in such a restrictive equation is impressive. This finding substantially reinforces the previously limited evidence that farm size gains influenced income.

The best and second-best R_3 equations both assign one variable to each of the three leading factors. The best R_3 equation is simply the best R_2 equation with SLxPOP%METRO (farm-urban mix) added. The second best R_3 equation (*a*) keeps MD-YRS-SCHOOL as the education variable but (*b*) changes the farm-urban variable from SLxPOP%METRO to POP%FARM and (*c*) changes the employment variable from POP%EMPLOYED to POP%LABOR, or the labor force participation rate. Note that employment, which ranked second in both of the R_2 equations, ranks first in the best R_3 equation and third in the second-best R_3 equation. So far, no factor has a clear advantage.

Four-Variable and Five-Variable Equations

Table 6.7 shows the best four-variable equation and three five-variable equations—best, second-best, and alternate. The R_4 equation starts with the best R_3 equation and adds a second education variable, COLLEGE%^{2}. The best R_5 equation (*a*) replaces the weaker education variable (COLLEGE%^{2}) with a second

Table 6.6
State/Nation Model: Best and Second-Best Two-Variable (R_2) and Three-Variable (R_3) Equations

	Best R_2		Second-Best R_2		Best R_3		Second-Best R_3	
	Parameter	t	Parameter	t	Parameter	t	Parameter	t
Constant	0.015	1.05	-0.075	-3.57	-0.001	-0.10	0.036	2.59
POP%FARM							-0.059	-5.54
FARM%xACRES			-0.578	-5.00				
SLxPOP%METRO					0.436	4.20		
MD-YRS-SCHOOL	0.877	6.40			0.679	5.39	0.847	6.60
POP%EMPLOYED	0.868	5.25	0.799	4.38	0.907	6.41		
POP%LABOR							1.321	5.17
R	.761		.695		.836		.808	
Adjusted R^2	.560		.460		.679		.620	

farm–urban variable, NSxPOP%FARM, and (*b*) replaces POP%EMPLOYED with the combination of POP%LABOR and UNEMPLOYED%; the equation is still limited to three factors.

The interesting thing about this equation is that farm–urban mix has an urban variable for the southern (SL) states and a farm variable for the nonsouthern (NS) states. Urban population rise and farm population decline are so tightly intertwined that we must be cautious about exaggerating the distinction. But the equation does suggest that the urban facet of farm–urban mix (higher incomes for urban populations) was most important in the South, whereas the agricultural facet (low incomes among farmers) was most important in the North. This suggestion is repeated in the second-best R_5 equation, which uses NSxPOP%TENANT in place of NSxPOP%FARM. Note that this equation is simply the best R_4 equation with NSxPOP%TENANT added.

We present the alternate equation, WAGE R_5, to demonstrate again that the wage rate is a significant influence but gets duplicated into oblivion. This equation is not specified according to objective criteria (best R and second-best R), so it does not contradict the evidence that none but the three strongest factors can get into the best or second-best equations of five variables or less. But the WAGE R_5 equation does show that WAGE-RATE (increase in average hourly earnings in manufacturing) is significant when only four other variables are controlled. One reason WAGE-RATE does all right in this equation is that MD-YRS-SCHOOL is not among the variables controlled. WAGE-RATE's intercorrelation with MD-YRS-SCHOOL is +.57, higher than WAGE-RATE's intercorrelation with any other variable. WAGE-RATE's t of 2.82 gains added stature when we recognize that WAGE-RATE also has very significant intercorrelations with two of the

Table 6.7
State/Nation Model: Best R_4 and Best, Second-Best, and Alternate R_5 Equations

	Best R_4		Best R_5		Second-Best R_5		WAGE R_5	
	Parameter	t	Parameter	t	Parameter	t	Parameter	t
Constant	-0.103	-3.10	0.308	2.53	-0.112	-3.52	-0.148	-4.79
SLxPOP%METRO	0.424	4.50	0.467	5.16	0.402	4.46	0.473	5.14
NSxPOP%FARM			-0.046	-4.63				
NSxPOP%TENANT					-18.70	-2.28	-24.59	-2.92
MD-YRS-SCHOOL	0.525	4.24	0.908	7.83	0.456	3.74		
COLLEGE%^{2}	0.333	3.27			0.406	3.97	0.458	4.34
POP%EMPLOYED	0.864	6.71			0.735	5.42	0.556	3.88
POP%LABOR			1.139	5.31				
UNEMPLOYED%			-0.086	-3.31				
WAGE-RATE							0.397	2.82
R	.871		.889		.886		.872	
Adjusted R^2	.737		.766		.760		.731	

variables that *are* controlled: +.46 with COLLEGE%^{2} and +.44 with SLxPOP%METRO. In other words, educational gains and metropolitan population gains in the South overlap the wage gains, yet controlling the two unfriendly variables is not enough to reduce WAGE-RATE to insignificance.

Six-Variable Equations

Table 6.8 has four 6-variable equations: best, second-best, and two alternate. The impressive thing about the best R_6 equation is that, even with six variables, it has room for only the three leading factors—farm-urban mix, education, and employment. Also noteworthy is the presence of (ACRES/FARM)^2. This is the first time farm size gains have had a variable to themselves in an equation, as opposed to being combined with farm population losses in an interaction variable. This equation is the fifth in which some variant of ACRES/FARM has appeared. So the evidence that farm size gains helped produce income gains is now reasonably solid. (The positive sign means that the biggest gains in average acres per farm are associated with the biggest gains in income.)

Each of the next three R_6 equations devotes four or five of its variables to representing the three leading factors. And two of those variables always

Table 6.8
State/Nation Model: Best, Second-Best, and Alternate R_6 Equations

	Best R_6		Second-Best R_6		First Alternate R_6		Second Alternate R_6	
	Parameter	t	Parameter	t	Parameter	t	Parameter	t
Constant	-0.082	-1.68	0.039	2.80	-0.000	-0.00	-0.071	-1.95
SLxPOP%METRO	0.467	5.44	0.448	4.66	0.409	4.60	0.541	6.28
POP%FARM			-0.042	-4.70				
NSxPOP%FARM	-0.047	-5.01						
NSxPOP%TENANT					-16.03	-2.00	-22.24	-2.83
(ACRES/FARM)^2	0.002	2.37						
MD-YRS-SCHOOL	0.808	6.87	0.678	4.66	0.558	4.32		
COLLEGE%					0.349	3.96		
COLLEGE%^{2}							0.406	4.09
POP%EMPLOYED					0.715	5.31	0.624	4.76
POP%LABOR	1.119	5.49	1.116	5.29				
UNEMPLOYED%	-0.084	-3.41	-0.081	-3.14				
POP%WID-DIVOR$^{.5}$					-0.317	-2.37		
NSxPOP%BLACK							-0.133	-2.41
TRANSCOSTDROP			-0.00004	-2.00			-0.00006	-3.89
R	.903		.897		.893		.893	
Adjusted R^2	.789		.776		.767		.767	

represent farm–urban mix; it emerges as generally the strongest factor in the four R_6 equations. But the leading factors are not what now concern us: the remaining R_6 equations are useful primarily for the light they shed on secondary factors. The second-best R_6 equation includes TRANSCOSTDROP. TRANSCOSTDROP is thus the first and only variable not representing one of the three leading factors that is strong enough to appear in an objectively chosen equation of six variables or less. The next equation, First Alternate R_6, features a different secondary variable: POP%WID-DIVOR$^{.5}$. This equation, of course, is not objectively chosen, but it does have a degree of objectivity: its R of .893 is barely below the .897 R of the second-best R_6 equation. Both the strength of the equation and the previous appearances of POP%WID-DIVOR in Percentage Change equations support a conclusion: increases in the widowed-divorced percentage restrain income growth.

The last R_6 equation has two secondary variables: TRANSCOSTDROP and NSxPOP%BLACK. By now we have seen TRANSCOSTDROP in so many equations that there is no room for reservations about its significance. And if its t of -2.00 in the second-best R_6 equation reaches only the 3 percent significance level, that modest showing is compensated for by TRANSCOSTDROP's new t of -3.89. This t is far beyond the 1 percent level, 2.43. The last R_6 equation's other secondary variable, NSxPOP%BLACK, is a new variation on an old theme. The Percentage Change Model's equations included two with POP%BLACK and three with SLxPOP%BLACK. Since POP%BLACK describes black population percentage variation in all 48 states of the sample, that variable does affirm that black population gains outside the South inhibited income growth. But SLxPOP%BLACK suggests that racial mix changes were influential mainly in the South. NSXPOP%BLACK counters this suggestion by declaring that racial mix changes were influential mainly *outside* the South. And when we review the full complement of Percentage Change and State/Nation equations (including two not yet discussed), we see that SLxPOP%BLACK and NSxPOP%BLACK are each in three equations and POP%BLACK is in two others. It is safe to conclude that racial mix changes influenced income growth in both the North and the South—but in different ways. Black population *losses* stimulated income growth in the South, whereas black population *gains* restrained income growth elsewhere.

Seven-Variable and Nine-Variable Equations

Table 6.9 shows the best State/Nation equations for seven and nine variables. The nine-variable equation, which has three versions, is also the best overall equation. No eight-variable equation is shown, but the best eight-variable equation is the basic nine-variable one with the weakest variable—UNEMPLOYED%—removed.

The R_7 equation has variables representing four factors: farm–urban mix, education, employment, and transportation. As usual, the first three factors are strongest: each leading factor has two variables to transportation's one, and even the weaker variable of each pair for the first three factors is stronger than TRANSCOSTDROP. Here, as in the second-best R_6 equation, TRANSCOSTDROP is the only variable that is strong enough to enter the equation yet does not represent one of the three leading factors. Transportation is, in short, a solid performer. It and racial mix are the only factors seriously competing for the honor of being ranked fourth in importance.

The best overall (R_9) equation for the State/Nation model, like the best overall equation for the Percentage Change model, has (1) a basic version, (2) an augmented version that adds SLAVE, and (3) a consolidated version that combines the individual variables for farm-urban mix, education, and employment into three compound variables—FARM-URBAN, EDUCATION, and EMPLOYMENT. The general procedure for computing compound variables is explained at the beginning of this chapter; the specific formulas for the three State/Nation compound variables appear at the bottom of table 6.9.

Table 6.9
State/Nation Model: Best R_7 Equation and Three Versions of Best R_9 Equation

	Best R_7		Basic Best R_9		Best R_9 Plus SLAVE		Consolidated Best R_9	
	Parameter	t	Parameter	t	Parameter	t	Parameter	t
Constant	-0.017	-0.46	0.065	0.84	0.085	1.02	0.065	1.00
SLxPOP%METRO	0.494	5.39	0.466	5.20	0.466	5.17		
NSxPOP%TENANT			-17.74	-2.34	-19.14	-2.43		
NSxPOP%FARM	-0.047	-5.06						
* FARM-URBAN							1.000	6.66
MD-YRS-SCHOOL	0.636	4.25	0.453	2.76	0.518	2.74		
COLLEGE%^{2}	0.206	1.88	0.351	3.66	0.354	3.66		
* EDUCATION							1.000	5.70
POP%EMPLOYED			0.485	3.02	0.524	3.07		
POP%LABOR	0.894	3.82						
UNEMPLOYED%	-0.076	-3.03	-0.052	-1.74	-0.045	-1.42		
* EMPLOYMENT							1.000	6.20
NSxPOP%BLACK			-0.101	-1.85	-0.112	-1.96	-0.101	-1.95
POP%WID-DIVOR$^{.5}$			-0.228	-1.84	-0.249	-1.94	-0.228	-2.00
TRANSCOSTDROP	-3.242	-1.82	-3.877	-2.00	-3.628	-1.83	-3.877	-2.26
SLAVE					-0.024	-0.70		
R	.907		.914		.916		.914	
Adjusted R^2	.792		.797		.795		.812	

* Compound variable (explained in text). FARM-URBAN = 0.466 SLxPOP%METRO - 17.74 NSxPOP%TENANT; EDUCATION = 0.351 COLLEGE%^{2} + 0.453 MD-YRS-SCHOOL; EMPLOYMENT = 0.485 POP%EMPLOYED - 0.052 UNEMPLOYED%.

Note: TRANSCOSTDROP is actually TRANSCOSTDROP/100. This conversion was necessary to avoid a regression coefficient whose first four digits were all zero. The conversion does not affect R or TRANSCOSTDROP's t.

Except for two minor substitutions of one variable for a highly similar one, the R_9 equation is the R_7 equation with two additional variables, each representing an additional factor. The minor substitutions are NSxPOP%TENANT for NSxPOP%FARM and POP%EMPLOYED for POP%LABOR. Both substitutes have about the same substance as their R_7 counterparts. The two additional variables are NSxPOP%BLACK and POP%WID-DIVOR$^{.5}$. Both were in one of

the alternate R_6 equations. Although the t's for the new variables (-1.85 and -1.84) reach only the 4 percent significance level here, racial mix and widowed-divorced female variables have now been in quite a few equations; we cannot doubt that their significance is genuine. NSxPOP%BLACK and POP%WID-DIVOR[5] reiterate that changes in racial mix and in the population percentage of widowed and divorced women caused changes in income.

In the R_9 equation's augmented version (Best R_9 Plus SLAVE), adding SLAVE to the equation has no appreciable effect on the t values of the three regional interaction variables (SLxPOP%METRO, NSxPOP%TENANT, and NSxPOP%BLACK). In fact, the t's for two of those variables actually increase, and the t for the third declines by a minuscule 0.03. These results demonstrate that none of the three interaction variables is primarily—or even secondarily—a proxy for SLAVE; the interaction is real, and the substantive parts of the interaction variables are in control. Given the equation's extensive control of the things that caused rapid income growth in the South, we are not surprised at the finding that the interaction variables can be taken at face value.

The consolidated equation (1) combines SLxPOP%METRO and NSxPOP%TENANT into FARM-URBAN, (2) combines MD-YRS-SCHOOL and COLLEGE%[2] into EDUCATION, and (3) combines POP%EMPLOYED and UNEMPLOYED% into EMPLOYMENT. The other three variables in the basic equation remain as separate variables, giving the consolidated equation a total of six variables. As previously explained, the purpose of consolidating all variables for a given factor into one compound variable is to eliminate multicollinearity within factors: multicollinearity (duplication and overlap among variables) causes t to be understated. The compound variables rank as follows: (1) FARM-URBAN, t = 6.66, (2) EMPLOYMENT, t = 6.20, and (3) EDUCATION, t = 5.70. The ranks of the three secondary variables are (4) TRANSCOSTDROP, t = -2.26, (5) POP%WID-DIVOR, t = -2.00, and (6) NSxPOP%BLACK, t = -1.95.

Consensus of State/Nation Model Evidence

Under the "Consensus" subheading in our discussion of the Percentage Change model, we went into a fair amount of detail in justifying our ranking for the factors. This detail was needed to show how we determined consensus and to demonstrate that the ranks assigned to the factors have an objective basis (notwithstanding some judgment). Having explained our ranking procedure for the primary model, we will forgo the detail for the secondary model, the State/Nation model. Suffice it to say that the ranking procedure used for the State/Nation model is essentially the same as that used for the Percentage Change model.

In the Percentage Change model, farm-urban mix and education were the two strongest determinants of income change and were judged equal in strength. The evidence from the State/Nation model points to the same consensus: farm-urban mix and education are again the two strongest factors and are tied for first rank. Employment again ranks third. It does perform a little better in the State/Nation model, particularly in the best R_2 and best R_3 equations. But ultimately

employment fails to match the performance of the two top-ranked factors. Racial mix (change in the black population percentage) and transportation (decline in the cost of shipping manufactured goods to the West) once more rank fourth and fifth. And, continuing the consensus parallels between the two models, the widowed-divorced female population percentage once more ranks sixth. The widowed-divorced factor, incidentally, is the only factor beyond the first five that is represented in the State/Nation equations.

Seventh, eighth, and ninth ranks respectively go to hourly wage changes, industry mix (manufacturing) changes, and retiree migration. These rankings are based largely on the simple *r*'s, although the WAGE R_5 equation does provide additional evidence supporting the wage factor. (WAGE-RATE has a *t* of 2.82 in that equation, which controls two farm–urban variables, an education variable, and an employment variable.) Retiree migration should really not be ranked: it shows no evidence of significance in the State/Nation model.

SUMMARY AND TWO-MODEL CONSENSUS

The Percentage Change and State/Nation models provide identical importance rankings through the first six ranks: (1 and 2) farm–urban mix and education, tied, (3) employment, (4) racial mix, (5) transportation, and (6) the widowed-divorced female percentage. Under the Percentage Change model the last three ranks are (7) retiree migration, (8) industry mix, or manufacturing, and (9) hourly wage rates. Under the State/Nation model, the last three ranks are (7) hourly wage rates, (8) industry mix, and (9) retiree migration. The only difference between the two rankings is, then, that the retiree migration and wage factors trade places.

The Percentage Change model is the better, more sensitive model. It does a better job of detecting and measuring relationships between income change in the abstract and the abstract factors that determine income change. That is, the variables in the Percentage Change model's Best-*R* equations yield higher *R*'s than the State/Nation model's variables for any fixed number of variables. And the best overall Percentage Change equation (R_{10}) produces a considerably higher *R*, .943, than the best overall State/Nation equation (R_9), which reaches only .914. Moreover, only the Percentage Change model is sensitive enough to detect the retiree migration influence. That influence not only registers, it registers in four equations and two variables (OLD-INMIGRATN and AMENITIES). The Percentage Change model also produces generally higher simple *r*'s: the highest Percentage Change *r* (+.69 for SLxPOP%URBAN) is well above the highest State/Nation *r* (+.59, also for SLxPOP%URBAN). Because the Percentage Change model performs better, we defer to its findings for the last three ranks.

One more repetition of the caveat about interpreting the factors enumerated above is appropriate. The simplified labels "farm–urban mix," "education," "employment," and the like refer to 1950–87 (sometimes 1950–80) *changes*, not to base-year *levels*. To avoid constantly repeating the word *changes* and creating excessively long and complicated labels for variables, we have adopted the

abbreviated terminology. Avoid any temptation to take our factor names and variable labels too literally: we are not referring to levels.

NOTES

1. Leonard F. Wheat, *Labor Force Participation and Unemployment in American Counties* (Washington, D.C.: Economic Development Administration, 1982), chapter 13, pp. 13-19, 13-26; table 13-6. Also summarized in Leonard F. Wheat, "The Determinants of 1963-77 Regional Manufacturing Growth: Why the South and West Grow," *Journal of Regional Science* 26 (1986), p. 644, and table 1.

2. By way of further clarification, the compound variable procedure is analogous to but better than factor analysis. Factor analysis likewise combines several original variables into one, and its compound variables, called factors, likewise describe multifaceted factors. But factor analysis has two drawbacks—beyond mathematical complexity—not found in the procedure used here. First, far from eliminating duplication among variables, factor analysis often *creates* duplication: it may put the same variable in more than one factor—often three or even four. Our procedure puts each variable being consolidated in one and only one factor. Second, the so-called factors created by factor analysis are often imaginary and even illogical: they can be purely mathematical creations with no real unifying substance; they typically (if not necessarily) lack the constraint of human judgment, not to mention common sense; they sometimes defy interpretation or—just as bad—seduce analysts into making arbitrary or even silly interpretations. Our combinations of variables are, in contrast, both rational and rationally determined. For example, POP%FARM, NSxTENANT%xAC, and SLxPOP%URBAN obviously all describe facets of the farm-to-urban population shift; MD-YRS-SCHOOL and COLLEGE% both measure change in educational attainment.

7

FURTHER ANALYSIS, SUMMARY, AND CONCLUSIONS

State per-capita incomes have been converging during most of this century, especially since 1940. When per-capita income ("income") is expressed as a percentage of the national average, the below-average states have been climbing and the above-average states descending. This convergence is primarily a matter of one far-below-average region's—the South's—catching up with the rest of the country. Income's relative decline in all other regions is due in large measure to fast southern growth's causing the national average to grow faster than incomes in the average and above-average regions. The main causes of convergence, therefore, are the forces that have brought fast income growth to the South. These southern forces have been heavily influenced by the collapse of the South's dominant economic institution: sharecropping (the sharecropper-tenant farmer system of agriculture). Outside the South, an especially important secondary influence has been transportation cost declines that appreciably lowered the prices of manufactured goods sold in the West, the region with the highest 1950 income. Farm-to-urban population shifts in the Plains and black population growth in the Manufacturing Belt have also played roles.

In this chapter we (1) identify our invalid hypotheses and subhypotheses and inquire about why they were wrong, (2) analyze the relationships among the factors that caused the South's rapid income growth, and (3) summarize our conclusions about the determinants of 1950–87 income change.

INVALID HYPOTHESES AND SUBHYPOTHESES

In chapter 4 we enumerated ten general hypotheses and a larger number of subhypotheses designed to explain income growth. These hypotheses attribute changes in income to *changes* in (1) farm-urban population mix, (2) average educational attainment, (3) worker-nonworker mix, or per-capita employment, (4)

the black population percentage, (5) the population percentages of other high-income and low-income groups, (6) hourly wage rates, (7) transportation costs, specifically, the proportionate amount of transportation costs embodied in the price of manufactured goods sold in the West, (8) industry mix, with emphasis on per-capita employment in manufacturing, (9) energy prices, and (10) short-run economic conditions affecting only certain regions and states.

Nine of the ten general hypotheses proved valid. Hypothesis 9, energy prices, proved to be invalid. And hypothesis 10, though valid, had to be brought under hypothesis 3: long-run (1950–87) employment change and short-run change (1980s) have so much overlap that we found it impossible to keep them separate. Hypothesis 10's main variable, UNEMPLOY%1987, duplicates hypothesis 3's variables—POP%EMPLOYED, POP%LABOR, and UNEMPLOYED%—to such an extent that both types of variables can rarely coexist in the same equation. Despite the validity of eight of the nine remaining hypotheses, many subhypotheses failed the test. In this section we review the invalid general hypothesis and the invalid subhypotheses.

The energy hypothesis concerns the effects of post-1973 energy price increases and post-1981 energy price declines. Three variables—ENERGY74/POP, ENERGY82/POP, and ENERGY-CRISIS—test the hypothesis. The first two measure energy industry employment; the third is a dummy identifying the 11 states most affected by post-1981 price declines. None of these variables proved significant. The likely explanation is that 1981–87 energy price decreases canceled the 1973–81 increases, leaving no significant net effect. The motor fuel price index as a percentage of the overall CPI was 70.3 percent for 1973, was 100 percent for the 1982–1983–1984 base period, and was 70.6 percent for 1987—virtually the same as for 1973. A supplemental explanation is that the overall employment variables (e.g., POP%EMPLOYED) overlapped the energy employment variables enough that controlling the employment variables made the energy variables insignificant. In this regard, ENERGY82/POP has r's of +.35 with UNEMPLOYED% and +.41 with UNEMPLOY%1987. (All statistics are from the Percentage Change model unless we state otherwise.)

Under the farm-urban mix hypothesis, the idea that declines in the rural nonfarm population percentage (POP%RURNONFARM) raise income got no support. In fact, POP%RURNONFARM sometimes became significantly positive—the wrong sign—and entered equations where it did not belong. (We rejected all equations using wrong-sign variables, because spurious correlation was probable and a good theoretical cause-and-effect relationship was absent.) When the variables in the longest equations were controlled, however, POP%RURNONFARM became insignificant. Its unexpected failure to be significantly negative may reflect a growing tendency of wealthy and affluent metropolitan residents to move into rural areas beyond the suburbs. There is a +.47 correlation (r) between POP%RURNONFARM and POP%METRO. Also, new manufacturing has tended to locate in exurbia—territory close to central cities but beyond the suburbs. We thus find a +.23 r (barely short of the 5 percent level, +.24) relating POP%RURNONFARM to HIMID-MFG/POP. Still another

consideration is that, partly as a result of improved highways, many rural residents now commute to metropolitan places of employment and thus earn metropolitan incomes. We are suggesting, in other words, that rural nonfarm residents may no longer have below-average incomes.

Under hypothesis 5 (high-income and low-income groups other than blacks), the subhypotheses concerning the Hispanic and Indian population percentages (POP%SPAIN and POP%INDIAN) failed badly. The flaw in the Hispanic subhypothesis may be that Hispanics tend to congregate in metropolitan areas, where wages are relatively high. Also, high female labor force participation rates for Hispanics tend to bolster Hispanic per-capita employment and income. Indians, for their part, may constitute such a small percentage of total population in most states that their effect is negligible. Moreover, Indian percentages are fairly stable.

Our industry mix hypothesis involved not just manufacturing but services and mining. Mining includes both (*a*) coal, metal, and nonmetallic minerals mining and quarrying and (*b*) oil and gas extraction. Our main subhypothesis, dealing with manufacturing, won support, but the service and mining subhypotheses were losers. The unsuccessful variables were SERVICE/POP, HI-SERV/POP (high-wage service industries), OIL-GAS/POP, MINING/POP, OIL&MINE/POP, and PETROLBRL/POP. The failure of service employment change to have an effect is not surprising. Service industry wages are a little on the low side but probably too close to average levels to affect the average significantly. Incidentally, HI-SERVE/POP displayed negative signs, thereby debunking our idea that growth in some service industries might actually raise income.

The mining variables were based on the premise that high wages in the heavily unionized mining sector would create a linkage between mining employment increases and income increases. But we found no linkage. We think the reason is that mining areas have low labor force participation rates—low percentages of income earners in the population. Women don't work in mines, so mining economies have low female labor force participation rates; the low female rates bring down the overall participation rates. In his study of county labor force participation rates, Wheat found a highly significant r of -.24 between the overall participation rate and mining employment as a percentage of total employment.[1] When the 2,706 counties in that study were grouped into ten participation rate intervals, the mining employment percentages were 0.5 percent in the highest participation interval, 0.8 percent in the second highest, 3.1 percent in the second lowest, and 8.1 percent in the lowest. The mining employment percentage was the sixth most significant variable (and was negative) in the best ten-variable equation for predicting county labor force participation rates.

Our original hypothesis 10 looked at the effects of short-run booms and busts that were limited to certain states or regions. One of the variables was FARM-CRISIS, a dummy identifying the 12 states most affected by post-1981 farm price declines. FARM-CRISIS was not significant. The probable reason is that the long-run (1950–87) tendency of the farm states to have high growth rates for income overpowered the short-run (1981–87) tendency for these states to have low growth rates. That is, the 1950–87 farm-urban variables incorporated 1981–87 changes.

FACTOR OVERLAP IN THE SOUTH

Rejection of one hypothesis (energy price effects) and several subhypotheses leaves us with eight valid hypotheses. The hypotheses involve *factors* (general influences) that caused income change. For analytical convenience, we treat the causal agents from hypothesis 5's two valid subhypotheses as two separate factors: widowed-divorced females and retiree migrants. We thus identify nine factors that significantly influenced income.

The next section of this chapter explains and evaluates the individual factors and summarizes their importance rankings. But before moving on to that subject we need to consider an analytical problem affecting the importance rankings. The problem is overlap among the factors that had the most influence in the South. Our purpose in this section is to show that three factors—racial mix, manufacturing, and wages—probably had their explanatory power understated in the regression analyses. The understatement resulted because stronger factors—chiefly farm-urban mix and education—served as proxies for (i.e., duplicated) the understated ones.

Five factors, and to a limited extent a sixth, had their strongest effects in the South. The five chief factors are (1) farm-urban mix, (2) education, (3) racial mix, or the black percentage, (4) hourly wage rates, and (5) manufacturing. Variables representing these factors tend to have their extreme values—their most favorable changes—in the South. (Manufacturing growth rates are second-highest in the South and highest in the West.) Consequently, these variables tend to duplicate each other and to act as proxies for each other. When a variable in one of the five categories is controlled, variables in the other categories are partly spoken for (represented by proxy) and thus tend to decline in significance.

The employment factor also had some southern emphasis, but this emphasis was minor. Female labor force participation rates, hence per-capita employment, were high among the sharecroppers: the women worked in the fields. The Great Black Migration therefore created a tendency for per-capita employment to decline in the South. This tendency partly offset a somewhat stronger tendency: the tendency for southern per-capita employment to rise as a result of nonsharecropper farm employment decline and reductions in employment discrimination against urban blacks. The net effect of these opposing tendencies is a +.10 intercorrelation between POP%EMPLOYED and SLAVE (southern location). But if POP%BLACK (change in the black population percentage) is controlled, a partial r of +.21 links POP%EMPLOYED and SLAVE. Though still short of significance (the 5 percent level is +.24), this r does reveal a latent tendency for per-capita employment gains to be above average in the South.

The duplication among the main five overlapping factors is strongest in the simple correlation situation, where nothing is held constant. But the problem remains even in the longest equations. In these equations the strongest variables—those representing farm-urban mix, education, and sometimes racial mix—crowd out the weakest ones. The crowded-out variables represent wages, manufacturing, and sometimes racial mix. Compelling reasons exist for believing that the slighted variables are genuinely significant, not just nothing-held-constant proxies for the strongest ones. Yet the slighted variables get left out of most equations, because

farm-urban and education variables (and sometimes one or two of the others) speak for them. Two particularly strong proxy tendencies are the tendency for education to represent racial mix and the tendency for urbanization to represent manufacturing.

Evidence in table 7.1 illustrates the problem of overlap among factors. The table is a correlation matrix covering the dependent variable (INCOME/POP:%), the slave state or South dummy (SLAVE), and the best variables for urbanization (SLxPOP%URBAN), agriculture (POP%TENANTFARM), education (MD-YRS-SCHOOL), racial mix (POP%BLACK), wages (WAGE-RATE), and manufacturing (HIMID-MFG/POP and MFG/POP). We use SLxPOP%URBAN (regional interaction variable) rather than POP%URBAN for urbanization because this particular subfactor—urbanization—had extraordinary power in the South. In the South, sharecropping made the 1950 disparity between farm income and urban income much larger than the disparity in other regions; the income benefits of the farm-to-urban shift were not simply a matter of bigger shifts in the South but, rather, the result of *interaction* between (*a*) proportionately bigger population shifts and (*b*) bigger income gains from a shift of any given magnitude. This interaction effect can be seen by comparing the simple r of +.17 for POP%URBAN (insignificant) with that of +.69 for SLxPOP%URBAN (the highest r in the study). Unfortunately, the SLAVE multiplier strengthens the proxy characteristics of the interaction variable (SLxPOP%URBAN), because SLAVE (SL) is a proxy for everything that caused rapid income growth in the South.

The first thing to observe in table 7.1 is the intercorrelations between SLAVE and the other independent variables (first column). These r's measure the extent to which the variables have their highest values (positive r's) or lowest values (negative r's) in southern states, where SLAVE has its highest values (dummy = 1). Although HIMID-MFG/POP's r of +.38 is above the 1 percent level (±.33 for the one-tail test), the two manufacturing variables have relatively low r's. But all the other SLAVE r's are extremely significant; their absolute values range from .55 to .88. Or, if we disregard SLxPOP%URBAN's .88 on grounds that this interaction variable includes SLAVE, the absolute values of r range from .55 to .70. About half of the variation in both POP%BLACK and POP%TENANTFARM can be explained by southern location. The SLAVE r's declare that (*a*) urbanization, education, and wages had generally large increases in the South and (*b*) agriculture and the black percentage had generally large decreases in the South. The findings thus show a high degree of geographic overlap among the factors.

The next thing to observe in table 7.1 is the first row's correlations between the dependent variable (INCOME/POP:%) and the independent variables (column headings). HIMID-MFG/POP again has a lower r than the other explanatory variables. But the absolute values of those other r's fall in a narrow range: .52 to .69. The similarity of these r's does not prove that the variables share the same substance (i.e., that the variables overlap), but that similarity does illustrate the *effect* of the known overlap. Because each of the five main explanatory variables (columns 2 through 6) overlaps the others and thus is partly a proxy for the others, their r's do not greatly differ. The r of the last variable, HIMID-

Table 7.1
Variables Representing Factors Whose Influence Was Greatest in the South: Correlations with Income and Intercorrelations among the Variables (Percentage Change Model)

	SLAVE	SLx POP% URBAN	POP% TENANT FARM	MD-YRS-SCHOOL	POP% BLACK	WAGE-RATE	HIMID-MFG/POP
INCOME/POP:%	.57	.69	-.52	.62	-.61	.59	.41
SLAVE		.88	-.69	.66	-.70	.55	.38
SLxPOP%URBAN	.88		-.68	.68	-.85	.51	.52
POP%TENANTFARM	-.69	-.68		-.62	.58	-.46	-.42
MD-YRS-SCHOOL	.66	.68	-.62		-.60	.57	.42
POP%BLACK	-.70	-.85	.58	-.60		-.42	-.61
WAGE-RATE	.55	.51	-.46	.57	-.42		.12
HIMID-MFG/POP	.38	.52	-.42	.42	-.61	.12	
MFG/POP	.12	.30	-.23	.19	-.39	-.04	.83

MFG/POP, falls below the others partly because HIMID-MFG/POP has the weakest proxy tendencies—the weakest intercorrelations with SLAVE and the other variables. (A sister variable, MFG/POP, actually has a weaker intercorrelation with SLAVE of only +.12, because manufacturing growth was highest not in the South but in the West.)

Not surprisingly, SLAVE's simple *r* of +.57 is roughly the average—in fact, it *is* the average—of the absolute values of the *r*'s in the other six columns. The explanation: SLAVE is a composite proxy for all of the influences that gave fast income growth to the South—urban population gains, farm population losses, educational gains, black population losses, wage gains, and manufacturing gains.

The table's remaining *r*'s—intercorrelations among the seven explanatory variables, including MFG/POP—show how much overlap there is between specific pairs of variables. As you would expect, the *r* relating urban population gain to tenant farm population loss is very high, -.68. Tenant farm loss shows almost as strong an association with educational gain (-.62). Since urban residents are generally better educated than farmers—tenant farmers in particular—these associations are understandable. The link between farm-urban mix and education is not particularly troublesome in the regression analysis, because the farm-urban and education variables are strong enough to take care of themselves. That is, both influences force their way into just about every best and second-best *R* equation. And with both controlled, each ceases to be a proxy for the other.

Farm-urban mix and education do, of course, continue to be proxies for the remaining southern influences—in addition to speaking for themselves. SLxPOP%URBAN, which is in most of the Percentage Change model's

equations, has a remarkable -.85 *r* with POP%BLACK. POP%BLACK also has *r*'s of +.58 with POP%TENANTFARM, +.42 with POP%FARM (not shown), -.60 with MD-YRS-SCHOOL, and -.39 with COLLEGE%^{2} (not shown). These *r*'s describe the out-migration of millions of poorly educated black sharecroppers from the South—a phenomenon that simultaneously increased the South's urban population percentage, decreased the South's tenant farm and overall farm population percentages, and increased the South's average level of educational attainment. From the regression standpoint, the effect is to squeeze POP%BLACK out of most equations by allowing farm-urban and education variables to serve as proxies for race variables.

The wage influence, represented by WAGE-RATE, has the same problem. WAGE-RATE has *r*'s of +.51 with SLxPOP%URBAN, -.46 with POP%TENANTFARM, +.57 with MD-YRS-SCHOOL, +.46 with COLLEGE%^{2} (not shown), and -.42 with POP%BLACK (which is in some equations). None of those intercorrelated variables is strongly enough related to WAGE-RATE to, by itself, put WAGE-RATE out of business. But the cumulative effect of an equation's having three, four, or even five variables that partly duplicate WAGE-RATE is devastating. And that doesn't end the duplication. The strongest nonsouthern variable, TRANSCOSTDROP, has a -.47 *r* with WAGE-RATE. This *r* reflects the fact that the highest 1950 transportation costs, hence the biggest transport cost declines, were in the West. The West also had the biggest relative wage declines. The transport cost declines, by lowering the cost of living, brought about the wage declines. So TRANSCOSTDROP can represent wage declines in the West, the region with the biggest relative wage changes outside the South.

The problem of duplication and overlap continues with HIMID-MFG/POP, which measures per-capita manufacturing employment change in the high-wage and mid-wage manufacturing industries. This variable does not get into any equations, although MFG/POP does enter a few. Manufacturing's lackluster performance is partly a matter of its being one of the weaker influences. But duplication is also at work. HIMID-MFG/POP has *r*'s of +.52 with SLxPOP%URBAN, -.42 with POP%TENANTFARM, +.42 with MD-YRS-SCHOOL, and -.61 with POP%BLACK. MFG/POP's intercorrelations with the same variables are much lower. (Compare the last two lines of table 7.1.) That is apparently why MFG/POP gets into some equations whereas HIMID-MFG/POP does not, although their respective simple *r*'s are +.15 and +.41. MFG/POP is *not* duplicated. MFG/POP's weaker intercorrelations with the southern-emphasis variables also help explain why MFG/POP's simple *r* is so much lower than HIMID-MFG/POP's. MFG/POP has less proxy power than HIMID-MFG/POP; HIMID-MFG/POP does not speak mainly for itself when it declares its simple *r* of +.41.

THE FACTORS THAT CAUSED CONVERGENCE

We have used correlation-regression analysis, evidence from other sources, knowledge of economic history, knowledge of social trends, and highly plausible

causal mechanisms (theory) to identify the factors responsible for the 1950–87 changes in relative income. Nine such factors have surfaced. All but one (widowed-divorced females) not only caused change but also contributed to income convergence; the exceptional factor may not have influenced convergence. Easily the most powerful factors were (1) gains in educational attainment, (2) farm-to-urban population shifts, and (3) increases in per-capita employment. The first two influences had so much force that, in the best two-variable (R_2) equation, one education and one farm–urban variable can jointly explain five-eighths of the variation in income change. Just three variables—two education, one farm–urban—can explain seven-tenths of the variance (best R_3 equation). Adding employment to the blend creates a trio of factors with prodigious explanatory power: the best six-variable (R_6) equation uses three farm–urban variables, two education variables, and one employment variable to explain more than five-sixths of the income variance.

The strongest secondary factors were (4) the migration of millions of blacks—poor people, mostly former sharecroppers—from the South to other regions and (5) the West's transportation cost decline, which led to a lower western price-wage structure, hence to lower relative income. Besides directly affecting income, the black migration contributed to the first three factors enumerated above and to the wage change factor, listed below. The remaining factors were (6) increases in the population percentage of widowed and divorced females, (7) the migration of affluent retirees, who tended to go to low-income states, (8) changes in manufacturing employment per capita, and (9) increases in hourly wage rates. When ranked by importance, these influences come in the order listed, except that the first two—education and farm-urban mix—are equally important.

Education

Changes in educational attainment profoundly affected per-capita income. This is understandable. A good education opens doors to better-paying jobs, makes a person more employable, and leads to the kind of employment that is relatively secure from layoffs and intermittent unemployment. College-educated persons usually enjoy professional employment, a source of above-average or even exceptionally high income.

Educational gains affected income in all states. But the greatest effect was in the South. The reason: sharecropping collapsed. Millions of former sharecroppers, mostly black and virtually all poor, left the South. By transferring poor and poorly educated people from the South to other regions, this migration raised the South's per-capita income and its average educational attainment while retarding gains in other regions. At the same time, other southern developments such as school desegregation brought better education to those who stayed behind. And spurred by southern industrialization (and perhaps encouraged by the arrival of air conditioning), educated whites from the North began moving into the South. Between 1950 and 1980, median years of schooling in the Deep South shot up from 8.0 to 12.2, or from 1.6 years below the Northeast's level to 0.3 year below. The education gap almost disappeared.

These developments register in the correlation-regression findings. Here we limit ourselves to evidence from the strongest model, the Percentage Change model. MD-YRS-SCHOOL has a higher simple r (+.62) than any other variable except SLxPOP%URBAN. And SLxPOP%URBAN must be discounted when nothing is controlled, because its dummy multiplier gives it exceptionally strong proxy support. Table 7.1 shows that SLxPOP%URBAN has higher intercorrelations with the other leading variables than MD-YRS-SCHOOL has, except in one instance. In the best equations for two, three, four, and five variables, education is the strongest factor. Education is also the strongest factor—by a thin margin—in the best overall equation, which has ten variables. All best-R equations of three or more variables have not just one but two education variables.

Farm-Urban Mix

Farm-to-urban and farm-to-metropolitan population shifts also affected per-capita income profoundly. (Some farms are in metropolitan areas, of course, but we are referring to shifts to the urban parts of metropolitan areas.) The reason for the income effects is that urban residents generally earn considerably more than farmers and farm laborers. The farm-urban income disparity is largest when the urban residents reside in metropolitan areas. Metropolitan urban residents have higher average incomes than other urban residents because metropolitan areas have (1) relatively high proportions of professional, managerial, technical, and skilled-trades employment, (2) relatively high costs of living—this results from higher housing, commercial rental, governmental (tax), and transportation costs—that produce wage-salary adjustments, and (3) relatively high proportions of employed women and teenagers.

As with education, the greatest effect was in the South. Because most southern farmers were sharecroppers and other tenant farmers with pitifully low incomes, southern farm incomes were far below nonsouthern farm incomes. Hence southern farm-to-urban shifts had more effect than northern shifts. At the same time, the South had bigger shifts—bigger declines in farm population. The South's population shifts were bigger because sharecropping collapsed, removing most of the sharecropper-tenants from the South's farm population. Though nonsouthern regions also lost farm population, their farms were much less labor-intensive in 1950: the nonsouthern regions had less change from labor-intensive to capital-intensive farming.

The correlation-regression findings reveal the effects of changes in farm-urban mix. SLxPOP%URBAN has the highest simple r (+.69) in the study; other impressive r's come from SLxPOP%METRO (+.61), SLxPOP%TENANT (-.58), POP%FARM (-.53), POP%METRO (+.52), and POP%TENANTFARM (-.52). Farm-urban mix is the strongest factor in several equations and, unlike education, never falls below second rank. In the most restrictive (best R_2) and best overall (R_{10}) equations, farm-urban mix is barely less significant than education. The best R_6 and R_{10} equations include three farm-urban variables apiece.

Employment

The third-ranking factor causing income change is employment. Employment is essentially per-capita employment, but the factor has labor force participation rate and unemployment facets too. The relationship between per-capita employment and per-capita income is obvious. Employed people earn income, whereas most people who are not employed, particularly those outside the labor force, do not earn income. High per-capita employment therefore brings high per-capita income. For about the same reason, high per-capita labor force participation brings high per-capita income. The labor force includes unemployed people who are seeking work in addition to employed people; unemployed people typically have some income (unemployment insurance). But because the unemployed have low income—sometimes none—high unemployment rates bring low per-capita income.

Employment gains too had more effect in the South than in most other regions. The analysis of factor overlap in the South revealed a latent relationship between per-capita employment (POP%EMPLOYED) and southern location (SLAVE). The South's per-capita employment (E/P) gains were relatively strong because

- educational gains in the South made southern blacks—and poor whites—more employable,
- southern employment discrimination against blacks declined,
- black percentages in nonsouthern states increased, so that discrimination in those states affected larger population percentages, and
- economic stagnation in the Manufacturing Belt caused North-to-South migration of working-age whites, reducing northern E/P relative to southern E/P.

The employment influence reveals its strength in POP%EMPLOYED's simple r: +.40. This is safely above the 1 percent level, ±.33. And when the strongest farm–urban variable, SLxPOP%URBAN, is controlled, POP%EMPLOYED climbs from +.40 to a partial r of +.53—high enough to explain one-fourth of the residual variance. Employment is represented in the second-best R_2 and R_3 equations, where it outperforms education, and it is in every best or second-best equation of four or more variables. In the best overall (R_{10}) equation and most others, employment is the third strongest factor.

Black Migration

Fourth in importance among the factors causing income change is black migration from the South to other regions, chiefly the Manufacturing Belt. This migration—Lemann calls it the Great Black Migration—was another outgrowth of the collapse of sharecropping.[2] Out of work, frequently ejected from their homes (i.e., shacks), and intrigued by reports of unheard-of wages being paid to

blacks in the North, millions of former sharecroppers flooded into northern cities. The migrants were almost always poor, sometimes destitute, so their movement affected income at both ends of the line. The population percentages of poor people went down in the South and up in the other regions. This geographical transfer of poverty naturally affected per-capita income.

Once again, the effect was greatest in the South. Black in-migration balanced black out-migration in the Peripheral South (six states), so the significant out-migration was largely confined to the Deep South (eight states). But the migrants distributed themselves over a much larger number of nonsouthern states, even if we disregard states where few blacks went. Consequently, the migrants constituted larger percentages of total population in states that gave than in states that received. Large population bases in the receiving states further diminished the in-migration percentages. Black migration's southern emphasis can also be viewed in terms of farm-urban mix and education. The out-migration directly reduced farm population percentages in the South, indirectly increased urban population percentages, and directly increased the population's average level of education.

The correlation-regression evidence affirms the importance of racial mix. POP%BLACK's simple *r* of +.61, barely lower than MD-YRS-SCHOOL's +.62, marks black migration as the third strongest influence by correlation test. The several black variables—POP%BLACK, SLxPOP%BLACK, and NSxPOP%BLACK—are much weaker in the regression tests, because farm-urban and education variables duplicate the race variables. But in the Percentage Change model, SLxPOP%BLACK identifies black migration as the fourth-strongest factor in two best or second-best equations and as the fifth-strongest factor in another. In the State/Nation model, NSxPOP%BLACK is in the best overall equation.

Western Transportation Cost Decline

Western transportation cost decline ranks fifth and is the second of the two strongest secondary influences. The prices of manufactured goods sold in the West declined relatively during 1950-87, because the transportation costs embodied in those prices declined. The transport cost declines were of two sorts. First, the cost of shipping goods long distances from the densely industrialized Manufacturing Belt to the thinly industrialized West declined. Improvements in highway and rail transportation—for example, the new interstate highway system—were responsible. Second, fewer goods had to be shipped those long distances, because the West began making more of its own goods. Throughout the period, the West had a higher regional manufacturing growth rate than any other region. The relative price reductions on manufactured goods lowered the West's relative cost of living. That effect, in turn, led to downward wage-salary adjustments. Average hourly wages in manufacturing for the West went from 114 percent of the national average in 1950 to 103 percent in 1987. (The only other region with a relative wage reduction was the Plains, where wages declined only from 98 percent to 97 percent.) The West's reduction in relative wages translates into relatively lower per-capita income.

Transportation is the only one of the five very strong or moderately strong

influences that lacks southern emphasis. This influence is a western influence, pure and simple. Theoretically, of course, the same developments could have affected both the South and the Plains, if to a lesser degree. But those regions are much closer to the Manufacturing Belt, hence did not have a long-distance shipping cost problem. Relative wages went up rather than down in the South and changed hardly at all in the Plains.

The transportation influence is strong enough to give TRANSCOSTDROP an r of -.36, three points above the 1 percent significance level. The related r^2, though not especially high, is strong enough to explain 13 percent of the variation in income. TRANSCOSTDROP is in many Percentage Change model equations. It performs well enough to make transportation the fourth strongest factor in the R_9 and R_{10} (best overall) equations. Transportation also ranks fourth in the State/Nation model's best R_7 and best R_9 (best overall) equations.

The Widowed-Divorced Percentage

The population percentage of widowed and divorced women ranks sixth in importance among the factors that caused income change. Both subgroups—widows and divorced women—have been growing steadily in proportion to total population. And the members of both subgroups tend to have below-average income. Differences in state growth rates for the widowed-divorced female percentage can thus influence income.

A rock-bottom simple r of -.01 for POP%WID-DIVOR invites pessimism about whether the variable will ultimately prove significant. This r also makes it apparent that the factor represented is not among the strongest. But when stronger influences are controlled, POP%WID-DIVOR does become significant. It is in the Percentage Change model's best R_8, R_9, and R_{10} equations. In these equations, the widowed-divorced factor respectively ranks seventh, sixth, and fifth; it averages sixth. In the State/Nation model, the only best-R equation using POP%WID-DIVOR is the best overall (R_9) equation. There the widowed-divorced factor ranks fifth among six factors.

Migration of Affluent Retirees

The seventh-ranked factor is retiree migration. America's population, including its retiree population, has become increasingly mobile in the decades since World War II. Many of the more affluent retirees move to states with amenity attractions—warm climate, seacoasts, lakes, and mountains. The destination states tend to be ones with average or below-average income—the Carolinas, Florida, Arkansas, Texas, New Mexico, and Arizona. By increasing the population percentages of affluent people in the destination states, the retiree migrants raise state income.

The chief migration variable, OLD-INMIGRATN, has an extremely insignificant simple r of +.06; a companion variable, AMENITIES, has a similar r of +.05. But when many other variables are controlled, both variables display

significance. OLD-INMIGRATN is in the Percentage Change model's second-best R_8, best R_9 and best overall (R_{10}) equations. Here the retiree factor respectively ranks fifth, fifth, and sixth. OLD-INMIGRATN is not in any of the State/Nation model's equations.

Manufacturing Change

Eighth in importance is manufacturing change, or growth and decline in manufacturing employment per capita. In general, manufacturing workers have above-average income. Adding people with above-average income to a state's population raises average income. And an indirect effect—urbanization—reinforces this direct effect. Powerful *r*'s of +.81 and +.72 relate MFG/POP and HIMID-MFG/POP to POP%URBAN. New manufacturing creates a first increment of population growth, which leads to increased purchases of goods and services, which creates a second (derived) increment of employment and population growth, which in turn has tertiary effects. The outcome is urban growth—urbanization. Urbanization brings improvements in occupational mix: higher proportions of professional, technical, managerial, and skilled trades employment. Higher income results.

Manufacturing change is another factor whose main effects are in the South. The South had high manufacturing growth rates throughout the 1950–87 period. The West, it is true, had even higher manufacturing growth rates. But, as we have seen, manufacturing growth in the West had a different effect: the transportation effect (lower prices on manufactured goods, a lower cost of living, lower wages and salaries, and consequently lower income). Because of the western offset, the simple *r*'s for manufacturing are on the low side: +.15 (insignificant) for MFG/POP and +.41 for HIMID-MFG/POP. And in the regression equations, the urbanization variables duplicate—indirectly control—manufacturing change. But MFG/POP is in the Percentage Change model's best R_8 equation. There the manufacturing factor actually ranks fourth in significance, right behind the three dominant factors.

Hourly Wage Rates

Ninth and last among the ranked factors is changes in hourly wage rates. Here we use the manufacturing wage rate as a proxy for the overall wage structure; using a specific one-digit sector—manufacturing—limits the influence of changes in industry mix on changes in relative wage levels. Hourly wage rates are the main determinant of annual income for a large segment of the population, so wage rate changes necessarily affect per-capita income.

The wage factor is yet another influence with a southern emphasis. Sharecropping had long been a source of surplus labor, a source that could be tapped to quench any incipient labor shortages in the South. (Sharecropper labor would, of course, move in at the bottom rung of the urban employment ladder, replacing other labor that moved up the ladder.) When sharecropping collapsed, the

sharecroppers either out-migrated or were slowly absorbed into the booming southern economy. By around 1970, the surplus labor was largely gone, and sharecropping no longer existed as a source of surplus labor. The South's continuing industrial growth therefore caused wages to climb. Wages in the Deep South moved up from 80 percent of the national average in 1950 to 87 percent in 1987.

The wage change influence registers in a hefty +.59 simple *r* for WAGE-RATE. This *r* is barely below the leading *r*'s of +.69 for SLxPOP%URBAN, +.62 for MD-YRS-SCHOOL, and -.61 for POP%BLACK. But as we have seen, WAGE-RATE is progressively duplicated by a host of farm-urban and education variables, and sometimes by a race variable too. WAGE-RATE is also duplicated by TRANSCOSTDROP, which indirectly describes declines in relative wage rates in the West. Consequently, WAGE-RATE is not in any of the best or second-best equations for given numbers of variables. But it is in some alternate equations that don't control MD-YRS-SCHOOL, and we see no reason to doubt that the wage influence is genuinely significant. In fact, if we acknowledge that the transportation influence could be reconceptualized as a wage influence, we can make a strong argument that the wage factor should be ranked fifth.

The Collapse of Sharecropping

A broader factor, the collapse of sharecropping in the South, contributed more to income convergence than did any of the nine comparatively narrow factors ranked above. Sharecropping's collapse cuts across three of the strongest factors (education, farm-urban mix, and racial mix) as well as hourly wages. For this reason, we have not ranked sharecropping as a separate factor. But the plain truth is that sharecropping was the foundation on which southern poverty rested. Sharecropping was what set the South apart from the rest of the country as a far-below-average income region. The South's low income was, more than anything else, a reflection of the huge class of agricultural workers—sharecroppers and other tenant farmers—whose abject poverty dragged southern income to its figurative knees. Since the South was the nation's most agricultural region, sharecropping's effect on southern income was all the stronger.

Sharecropping's heyday was the era that ended around 1950. Then came the mechanical cotton picker, mechanized cultivation, herbicides, a shift from cotton to less labor-intensive crops, mechanized tobacco harvesting, and other developments. These brought the sharecropper-tenant farmer system to a rather abrupt end. There were many income repercussions. Sharecropping's demise

- drastically shrank the South's farm population percentage,
- left the South with a more prosperous class of farmers,
- gave the South a more urbanized society, one with a better occupational mix and higher per-capita employment,

- helped raise the South's average level of educational attainment (because millions of poorly educated sharecroppers left the South),
- lowered the South's black population percentage while raising the North's, and
- eliminated the source of the surplus labor that was holding southern wages down.

The overall effect of these overlapping changes was striking improvement in the South's income position. This improvement was the essence of the convergence of state per-capita incomes that occurred during 1950–87.

NOTES

1. Leonard F. Wheat, *Labor Force Participation and Unemployment in American Counties* (Washington, D.C.: Economic Development Administration, 1982), pp. 8–22, 8–23; tables 3–2, 4–3.

2. Nicholas Lemann, *The Promised Land: The Great Black Migration and How It Changed America* (New York: Knopf, 1991).

BIBLIOGRAPHY

Alston, Lee J., and Joseph P. Ferrie. "Paternalism in Agricultural Labor Contracts in the U.S. South: Implications for the Growth of the Welfare State." *American Economic Review* 83 (1993), 852-76.

Amos, Orley M., Jr. "An Inquiry into the Causes of Increasing Regional Income Inequality in the United States." *Review of Regional Studies* 19 (Spring 1989), 1-12.

Barro, Robert J., and Xavier Sala-i-Martin. "Convergence." *Journal of Political Economy* 100 (April 1992), 223-51.

Bass, Bernard M., and Ralph Alexander. "Climate, Economy, and the Differential Migration of White and Nonwhite Workers." *Journal of Applied Psychology* 56 (1972), 518-21.

Blanchard, Oliver J., and Lawrence F. Katz. "Regional Evolutions." *Brookings Papers on Economic Activity* 1 (1992), 1-74.

Browne, Lynn E. "The Shifting Pattern of Interregional Migration." *New England Economic Review* (November-December 1979), 17-32.

———. "Narrowing Regional Income Differentials: I." *New England Economic Review* (September-October 1980), 35-54.

———. "Narrowing Regional Income Differentials: II." *New England Economic Review* (November-December 1980), 40-57.

———. "Shifting Regional Fortunes: The Wheel Turns." *New England Economic Review* (May-June 1989), 27-46.

Chinitz, Benjamin. "The Regional Transformation of the American Economy." *Urban Studies* 23 (1986), 377-85.

———, and Raymond Vernon. "Changing Forces in Industrial Location." *Harvard Business Review* 38 (1960), 126-36.

Connaughton, John E., and Ronald A. Madsen. "The Changing Regional Structure of the U.S. Economy." *Growth and Change* 23 (1990), 48-60.

Coughlin, Cletus C., and Thomas B. Mandelbaum. "Why Have State Per Capita Incomes Diverged Recently?" *Federal Reserve Bank of St. Louis Review* 70 (1988), 24-30.

Crandall, Robert W. *Manufacturing on the Move*. Washington, D.C.: Brookings Institution, 1993.

Crown, William H. "State Economic Implications of Elderly Interstate Migration." *The Gerontologist* 28 (1988), 533–39.

———. *Manufacturing Employment Growth in Urbanized Nonmetropolitan Counties.* Washington, D.C.: Economic Development Administration, 1989.

Crown, William H., and Stephen F. Fournier. *Migration and Regional Development.* Washington, D.C.: Economic Development Administration, 1987.

Crown, William H., Phyllis H. Mutschler, James H. Schulz, and Rebecca Loew. 1993. *The Economic Status of Divorced Older Women.* Waltham, Mass.: Policy Center on Aging, Brandeis University.

Cushing, Brian J. "Location-Specific Amenities, Topography, and Population Migration." *Annals of Regional Science* 21 (1987), 74–85.

Day, Richard H. "The Economics of Technological Change and the Demise of the Sharecropper." *American Economic Review* 57 (1967), 427–49.

Duffy, Neal E. "The Determinants of State Manufacturing Growth Rates: A Two-Digit-Level Analysis." *Journal of Regional Science* 34 (1994), 137–62.

Fantus Company, Inc. *The Appalachian Location Research Studies Program: Summary Report and Recommendations.* Washington, D.C.: Appalachian Regional Commission, 1966.

Fuchs, Victor R. *Changes in Location of Manufacturing Since 1929.* New Haven: Yale University Press, 1962.

Galbraith, John Kenneth. *American Capitalism: The Concept of Countervailing Power.* Boston: Houghton Mifflin, 1952.

Garnick, Daniel H. "Accounting for Regional Differences in Per Capita Personal Income Growth: An Update and Extension." *Survey of Current Business* 70 (January 1990), 29–40.

Garnick, Daniel H., and Howard L. Friedenberg. "Accounting for Regional Differences in Per Capita Personal Income Growth, 1929–79." *Survey of Current Business* 62 (September 1982), 24–34.

Ginzberg, Eli. "Youth Unemployment." *Scientific American* 242 (May 1980), 43–49.

Goode, William. *The Family.* Englewood Cliffs, N.J.: Prentice-Hall, 1982.

Graves, Philip E. "Migration and Climate." *Journal of Regional Science* 20 (1980), 227–37.

Greenhut, Melvin L. "An Empirical Model and a Survey: New Plant Locations in Florida." *Review of Economics and Statistics* 41 (1959), 433–38.

———. "An Explanation of Industrial Development in Underdeveloped Areas of the United States." *Land Economics* 36 (1960), 371–79.

Haas, W., and W. Serow. *The Influence of Retirement In-Migration on Local Economic Development: Final Report to the Appalachian Regional Commission.* Asheville, N.C.: University of North Carolina, 1990.

Hanna, Frank A. *State Income Differentials, 1919–1954.* Durham: Duke University Press, 1959.

Hansen, Niles. "Economic Development and Regional Heterogeneity: A Reconsideration of Regional Policy for the United States." *Economic Development Quarterly* 2 (1988), 107–18.

Heckscher, Eli. "The Effect of Foreign Trade on the Distribution of Income." *Ekonomisk Tidskrift* 21 (1919), 497–512.

Johnson, Kenneth P., John R. Kort, and Howard L. Friedenberg. "Regional and State Projections of Income, Employment, and Population to the Year 2000." *Survey of Current Business* 70, no. 5 (1990), 33–40.

Koshal, Rjindar K. "Urban Growth and Environment." *Growth and Change* 7 (1976), 34–38.

Lemann, Nicholas. "The Origins of the Underclass." *Atlantic Monthly* 259 (June 1986),

31–55.

———. *The Promised Land: The Great Black Migration and How It Changed America.* New York: Knopf, 1991.

Lichtenberg, Robert M. *One-Tenth of a Nation.* Cambridge: Harvard University Press, 1960.

Long, John F. *Population Deconcentration in the United States.* Washington, D.C.: Bureau of the Census, 1981.

Long, Larry H., and Kristin A. Hansen. *Reasons for Interstate Migration: Jobs, Retirement, Climate, and Other Influences.* Washington, D.C.: Bureau of the Census, 1979.

Lonsdale, Richard E., and H. L. Seyler. *Nonmetropolitan Industrialization.* New York: John Wiley & Sons, 1979.

McLaughlin, Glenn E., and Stefan Robock. *Why Industry Moves South.* Washington, D.C.: National Planning Association, 1949.

Maddox, James G., with E. E. Liebhafsky, Vivian W. Henderson, and Herbert M. Hamlin. *The Advancing South: Manpower Prospects and Problems.* New York: Twentieth Century Fund, 1967.

Mallick, Rajiv. "Convergence of State Per Capita Incomes: An Examination of Its Sources." *Growth and Change* 24 (1993), 321–40.

Marshall, Roy. *The State of Families.* Milwaukee: Family Service America, 1991.

Nissan, Edward, and George Carter. "Income Inequality across Regions over Time." *Growth and Change* 24 (1993), 303–19.

Norton, R. D. "Industrial Policy and American Renewal." *Journal of Economic Literature* 24 (1986), 1–40.

Ohlin, Bertil. *Interregional and International Trade.* Cambridge: Harvard University Press, 1933.

Olson, Mancur. *The Rise and Decline of Nations: Economic Growth, Stagflation, and Social Rigidities.* New Haven: Yale University Press, 1982.

Olvey, Lee D. "Regional Growth and Inter-Regional Migration—Their Pattern of Interaction." *Review of Regional Studies* 2 (1972), 139–63.

Perloff, Harvey S., Edgar S. Dunn, Jr., Eric E. Lampard, and Richard F. Muth. *Regions, Resources and Economic Growth.* Baltimore: Johns Hopkins University Press, 1960.

Persky, Joseph. "Push and Pull in Migration from Southern Farms." *Review of Regional Studies* 2 (1972), 49–62.

Rees, Albert. "An Essay on Youth Joblessness." *Journal of Economic Literature* 24 (1986), 613–28.

Rowley, Thomas D., John M. Redman, and John Angle. *The Rapid Rise in State Per Capita Income Inequality in the 1980s.* Washington, D.C.: U.S. Department of Agriculture, Economic Research Service, 1991.

Schmenner, Roger W. *The Manufacturing Location Decision: Evidence from Cincinnati and New England.* Washington, D.C.: Economic Development Administration, 1978.

———. *Making Business Location Decisions.* Englewood Cliffs, N.J.: Prentice-Hall, 1982.

Schwirian, Kent P. *Determinants of County and City Median Family Income Levels and Poverty Rates.* Washington, D.C.: Economic Development Administration, 1991.

Stafford, Howard A. *Principles of Industrial Facility Location.* Atlanta: Conway Publications, 1979.

———. "Environmental Protection and Industrial Location." *Annals of the Association of American Geographers* 75 (1985), 227–49.

Svarte, Larry M. "Environmental Preference Migration: A Review." *Geographical Review* 66 (1976), 314–30.

Thompson, Wilbur R., and John M. Mattila. *An Econometric Model of Postwar State Industrial Development.* Detroit: Wayne State University Press, 1959.

Truman, David B. *The Governmental Process: Political Interests and Public Opinion.* New York: Alfred A. Knopf, 1951.

Ullman, Edward L. "Amenities as a Factor in Regional Growth." *Geographical Review* 44 (1954), 119–32.

U.S. Department of Agriculture, Economic Research Service. *Rural People in the American Economy.* Washington, D.C.: 1966.

Wheat, Leonard F. *Regional Growth and Industrial Location: An Empirical Viewpoint.* Lexington, Mass.: D.C. Heath, 1973.

———. "Determinants of Median Family Income" (Appendix A). In *Labor Force Participation and Unemployment in American Counties.* Washington, D.C.: Economic Development Administration, 1982.

———. "The Determinants of 1963–77 Regional Manufacturing Growth: Why the South and West Grow." *Journal of Regional Science* 26 (1986), 635–59.

———. "The Trouble with Sclerosis: What Causes Regional Growth Differences?" Paper presented at the Regional Science Meetings, Boston, November 1990.

Wheat, Leonard F., John M. Wardwell, and Lee Faulkner. *The 1970s Migration Turnaround in Rural, Nonadjacent Counties.* Washington, D.C.: Economic Development Administration, 1984.

Wheeler, James O. "Effects of Geographical Scale on Location Decisions in Manufacturing: The Atlanta Example." *Economic Geography* 57 (1981), 134–45.

Williamson, Jeffrey G. "Regional Inequality and the Process of National Development: A Description of Patterns." *Economic Development and Cultural Change* 12 (July 1965), 3–45.

Wright, Gavin. *Old South, New South: Revolutions in the Southern Economy Since the Civil War.* New York: Basic Books, 1986.

INDEX

About the Authors

LEONARD F. WHEAT is an economist with the Economic Development Administration of the U.S. Department of Commerce.

WILLIAM H. CROWN is Associate Professor at Brandeis University.

Both have published extensively on topics in regional economics.